THE COLLECTED PAPERS OF ALBERT EINSTEIN

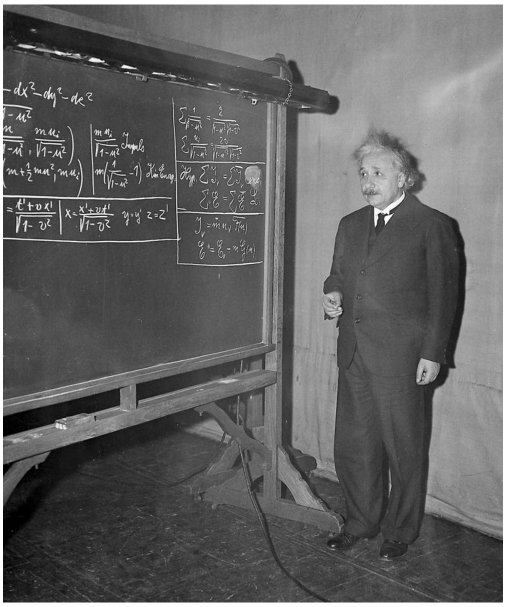
VOL. 6

ENGLISH TRANSLATIONS OF SELECTED TEXTS

The Foundation of the General Theory of Relativity

Author:
Albert EINSTEIN

(⋆ March, 14., 1879 in Ulm;
† April, 18., 1955 in Princeton, New Jersey)



The Berlin Years, 1914 – 1917

Abstract

The theory which is presented in the following pages conceivably constitutes the farthest-reaching generalization of a theory which, today, is generally called the "***theory of relativity***"; I will call the latter one – in order to distinguish it from the first named – the "special theory of relativity", which I assume to be known. The generalization of the theory of relativity has been facilitated considerably by MINKOWSKI, a mathematician who was the first one to recognize the formal equivalence of space coordinates and the time coordinate, and utilized this in the construction of the theory. The mathematical tools that are necessary for general relativity were readily available in the "absolute differential calculus", which is based upon the research on non-EUCLIDean manifolds by GAUSS, RIEMANN, and CHRISTOFFEL, and which has been systematized by RICCI and LEVI-CIVITA and has already been applied to problems of theoretical physics. In Part II of the present paper I developed all the necessary mathematical tools – which cannot be assumed to be known to every physicist – and I tried to do it in as simple and transparent a manner as possible, so that a special study of the mathematical literature is not required for the understanding of the present paper. Finally, I want to acknowledge gratefully my friend, the mathematician MARCEL GROSSMANN, whose help not only saved me the effort of studying the pertinent mathematical literature, but who also helped me in my search for the field equations of gravitation.

Contents

Part I.

FUNDAMENTAL CONSIDERATIONS ON THE POSTULATE OF RELATIVITY

1. Observations on the Special Theory of Relativity

The special theory of relativity is based on the following postulate, which is also satisfied by the mechanics of Galileo and Newton. If a system of co-ordinates K is chosen so that, in relation to it, physical laws hold good in their simplest form, the *same* laws also hold good in relation to any other system of co-ordinates K' moving in uniform translation relatively to K. This postulate we call the "special principle of relativity". The word "special" is meant to intimate that the principle is restricted to the case when K' has a motion of uniform translation relatively to K, but that the equivalence of K' and K does not extend to the case of non-uniform motion of K' relatively to K.
Thus the special theory of relativity does not depart from classical mechanics through the postulate of relativity, but through the postulate of the constancy of the velocity of light *in vacuo*, from which, in combination with the special principle of relativity, there follow, in the well-known way, the relativity of simultaneity, the Lorentzian transformation, and the related laws for the behaviour of moving bodies and clocks.
The modification to which the special theory of relativity has subjected the theory of space and time is indeed far reaching, but one important point has remained unaffected. For the laws of geometry, even according to the special theory of relativity, are to be interpreted directly as laws relating to the possible relative positions of solid bodies at rest; and, in a more general way, the laws of kinematics are to be interpreted as laws which describe the relations of measuring bodies and clocks. To two selected material points of a stationary rigid body there always corresponds a distance of quite definite length, which is independent of the locality and orientation of the body, and is also independent of the time. To two selected positions of the hands of a clock at rest relatively to the privileged system of reference there always corresponds an interval of time of a definite length, which is independent of place and time. We shall soon see that the general theory of relativity cannot adhere to this simple physical interpretation of space and time.

2. The Need for an Extension of the Postulate of Relativity

In classical mechanics, and no less in the special theory of relativity, there is an inherent epistemological defect which was, perhaps for the first time, clearly pointed out by Ernst Mach. We will elucidate it by the following example:– Two fluid bodies of the same size and nature hover freely in space at so great a distance from each other and from all other masses that only those gravitational forces need be taken into account which arise from the interaction of different parts of the same body. Let the distance between the two bodies be invariable, and in neither of the bodies let there be any relative movements of the parts with respect to one another. But let either mass, as judged by an observer at rest relatively to the other mass, rotate with constant angular velocity about the line joining the masses. This is a verifiable relative motion of the two bodies. Now let us imagine that each of the bodies has been surveyed by means of measuring instruments at rest relatively to itself, and let the surface of S_1 prove to be a sphere, and that of S_2 an ellipsoid of revolution. Thereupon we put the question – What is the reason for this difference in the two bodies? No answer can be admitted as epistemologically satisfactory,[1] unless the reason given is an *observable fact of experience*. The law of causality has not the significance of a statement as to the world of experience, except when *observable* facts ultimately appear as causes and effects.

Newtonian mechanics does not give a satisfactory answer to this question. It pronounces as follows:– The laws of mechanics apply to the space R_1, in respect to which the body S_1 is at rest, but not to the space R_2, in respect to which the body S_2 is at rest. But the privileged space R_1 of Galileo, thus introduced, is a merely *factitious* cause, and not a thing that can be observed. It is therefore clear that Newton's mechanics does not really satisfy the requirement of causality in the case under consideration, but only apparently does so, since it makes the factitious cause R_1 responsible for the observable difference in the bodies S_1 and S_2.

The only satisfactory answer must be that the physical system consisting of S_1 and S_2 reveals within itself no imaginable cause to which the differing behaviour of S_1 and S_2 can be referred. The cause must therefore lie *outside* this system. We have to take it that the general laws of motion, which in particular determine the shapes of S_1 and S_2, must be such that the mechanical behaviour of S_1 and S_2 is

[1] Of course an answer may be satisfactory from the point of view of epistemology, and yet be unsound physically, if it is in conflict with other experiences.

partly conditioned, in quite essential respects, by distant masses which we have not included in the system under consideration. These distant masses and their motions relative to S_1 and S_2 must then be regarded as the seat of the causes (which must be susceptible to observation) of the different behaviour of our two bodies S_1 and S_2. They take over the role of the factitious cause R_1. Of all imaginable spaces R_1, R_2, etc., in any kind of motion relatively to one another, there is none which we may look upon as privileged *a priori* without reviving the above-mentioned epistemological objection. *The laws of physics must be of such a nature that they apply to systems of reference in any kind of motion.* Along this road we arrive at an extension of the postulate of relativity.

In addition to this weighty argument from the theory of knowledge, there is a well-known physical fact which favours an extension of the theory of relativity. Let K be a GALILEAN system of reference, i.e. a system relatively to which (at least in the four-dimensional region under consideration) a mass, sufficiently distant from other masses, is moving with uniform motion in a straight line. Let K' be a second system of reference which is moving relatively to K in *uniformly accelerated* translation. Then, relatively to K', a mass sufficiently distant from other masses would have an accelerated motion such that its acceleration and direction of acceleration are independent of the material composition and physical state of the mass.

Does this permit an observer at rest relatively to K' to infer that he is on a "really" accelerated system of reference? The answer is in the negative; for the above-mentioned relation of freely movable masses to K' may be interpreted equally well in the following way. The system of reference K' is unaccelerated, but the space-time territory in question is under the sway of a gravitational field, which generates the accelerated motion of the bodies relatively to K'.

This view is made possible for us by the teaching of experience as to the existence of a field of force, namely, the gravitational field, which possesses the remarkable property of imparting the same acceleration to all bodies.[2] The mechanical behaviour of bodies relatively to K' is the same as presents itself to experience in the case of systems which we are wont to regard as "stationary" or as "privileged". Therefore, from the physical standpoint, the assumption readily suggests itself that the systems K and K' may both with equal right be looked upon as "stationary", that is to say, they have an equal title as systems of reference for the physical description of phenomena.

It will be seen from these reflexions that in pursuing the general theory of relativity we shall be led to a theory of gravitation, since we are able to "produce" a gravitational field merely by changing the system of co-ordinates. It will also be obvious that the principle of the constancy of the velocity of light *in vacuo* must be modified, since we easily recognize that the path of a ray of light with respect to K' must in general be curvilinear, if with respect to K light is propagated in a straight line with a definite constant velocity.

[2] EÖTVÖS has proved experimentally that the gravitational field has this property in great accuracy.

3. The Space-Time Continuum. Requirement of General Co-Variance for the Equations Expressing General Laws of Nature

In classical mechanics, as well as in the special theory of relativity, the co-ordinates of space and time have a direct physical meaning. To say that a point-event has the X_1 coordinate x_1 means that the projection of the point-event on the axis of X_1, determined by rigid rods and in accordance with the rules of EUCLIDean geometry, is obtained by measuring off a given rod (the unit of length) x_1 times from the origin of co-ordinates along the axis of X_1. To say that a point-event has the X_4 co-ordinate $x_4 = t$, means that a standard clock, made to measure time in a definite unit period, and which is stationary relatively to the system of co-ordinates and practically coincident in space with the point-event,[1] will have measured off $x_4 = t$ periods at the occurrence of the event.

This view of space and time has always been in the minds of physicists, even if, as a rule, they have been unconscious of it. This is clear from the part which these concepts play in physical measurements; it must also have underlain the reader's reflexions on the preceding chapter 2 for him to connect any meaning with what he there read. But we shall now show that we must put it aside and replace it by a more general view, in order to be able to carry through the postulate of general relativity, if the special theory of relativity applies to the special case of the absence of a gravitational field.

In a space which is free of gravitational fields we introduce a GALILEAN system of reference $K(x, y, z, t)$, and also a system of co-ordinates $K'(x', y', z', t')$ in uniform rotation relatively to K. Let the origins of both systems, as well as their axes of Z, permanently coincide. We shall show that for a spacetime measurement in the system K' the above definition of the physical meaning of lengths and times cannot be maintained. For reasons of symmetry it is clear that a circle around the origin

[1] We assume the possibility of verifying "simultaneity" for events immediately proximate in space, or – to speak more precisely – for immediate proximity or coincidence in space-time, without giving a definition of this fundamental concept.

in the X, Y plane of K may at the same time be regarded as a circle in the X', Y' plane of K'. We suppose that the circumference and diameter of this circle has been measured with a unit measure infinitely small compared with the radius, and that we have the quotient of the two results. If this experiment were performed with a measuring-rod at rest relatively to the GALILEAN system K, the quotient would be π. With a measuring-rod at rest relatively to K', the quotient would be greater than π. This is readily understood if we envisage the whole process of measuring from the "stationary" system K, and take into consideration that the measuring-rod applied to the periphery undergoes a LORENTZian contraction, while the one applied along the radius does not. Hence EUCLIDean geometry does not apply to K'. The notion of co-ordinates defined above, which presupposes the validity of EUCLIDean geometry, therefore breaks down in relation to the system K'. So, too, we are unable to introduce a time corresponding to physical requirements in K', indicated by clocks at rest relatively to K'. To convince ourselves of this impossibility, let us imagine two clocks of identical constitution placed, one at the origin of co-ordinates, and the other at the circumference of the circle, and both envisaged from the "stationary" system K. By a familiar result of the special theory of relativity, the clock at the circumference – judged from K – goes more slowly than the other, because the former is in motion and the latter at rest. An observer at the common origin of co-ordinates, capable of observing the clock at the circumference by means of light, would therefore see it lagging behind the clock beside him. As he will not make up his mind to let the velocity of light along the path in question depend explicitly on the time, he will interpret his observations as showing that the clock at the circumference "really" goes more slowly than the clock at the origin. So he will be obliged to define time in such a way that the rate of a clock depends upon where the clock may be.
We therefore reach this result:– In the general theory of relativity, space and time cannot be defined in such a way that differences of the spatial co-ordinates can be directly measured by the unit measuring-rod, or differences in the time co-ordinate by a standard clock.
The method hitherto employed for laying co-ordinates into the space-time continuum in a definite manner thus breaks down, and there seems to be no other way which would allow us to adapt systems of co-ordinates to the four-dimensional universe so that we might expect from their application a particularly simple formulation of the laws of nature. So there is nothing for it but to regard all imaginable systems of co-ordinates, on principle, as equally suitable for the description of nature. This comes to requiring that:–

> *The general laws of nature are to be expressed by equations which hold good for all systems of co-ordinates, that is, are co-variant with respect to any substitutions whatever (generally co-variant).*

It is clear that a physical theory which satisfies this postulate will also be suitable for the general postulate of relativity, for the sum of *all* substitutions in any case includes those which correspond to all relative motions of three-dimensional systems of co-ordinates. That this requirement of general co-variance, which takes away from space and time the last remnant of physical objectivity, is a natural one, will be seen from the following reflexion. All our space-time verifications invariably amount to a determination of space-time coincidences. If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meetings of two or more of these points. Moreover, the results of our measuring are nothing but verifications of such meetings of the material points of our measuring instruments with other material points, coincidences between the hands of a clock and points on the clock dial, and observed point-events happening at the same place at the same time.

The introduction of a system of reference serves no other purpose than to facilitate the description of the totality of such coincidences. We allot to the universe four space-time variables x_1, x_2, x_3, x_4 in such a way that for every point-event there is a corresponding system of values of the variables x_1, x_2, x_3, x_4. To two coincident point-events there corresponds one system of values of the variables x_1, x_2, x_3, x_4, i.e. coincidence is characterized by the identity of the co-ordinates. If, in place of the variables x_1, x_2, x_3, x_4, we introduce functions of them, x'_1, x'_2, x'_3, x'_4, as a new system of co-ordinates, so that the systems of values are made to correspond to one another without ambiguity, the equality of all four co-ordinates in the new system will also serve as an expression for the space-time coincidence of the two point-events. As all our physical experience can be ultimately reduced to such coincidences, there is no immediate reason for preferring certain systems of co-ordinates to others, that is to say, we arrive at the requirement of general co-variance.

4. The Relation of the Four Co-ordinates to Measurement in Space and Time

It is not my purpose in this discussion to represent the general theory of relativity as a system that is as simple and logical as possible, and with the minimum number of axioms; but my main object is to develop this theory in such a way that the reader will feel that the path we have entered upon is psychologically the natural one, and that the underlying assumptions will seem to have the highest possible degree of security. With this aim in view let it now be granted that:–
For infinitely small four-dimensional regions the theory of relativity in the restricted sense is appropriate, if the co-ordinates are suitably chosen. For this purpose we must choose the acceleration of the infinitely small (”local”) system of co-ordinates so that no gravitational field occurs; this is possible for an infinitely small region. Let X_1, X_2, X_3, be the co-ordinates of space, and X_4 the appertaining co-ordinate of time measured in the appropriate unit.[1] If a rigid rod is imagined to be given as the unit measure, the co-ordinates, with a given orientation of the system of co-ordinates, have a direct physical meaning in the sense of the special theory of relativity. By the special theory of relativity the expression

$$\mathrm{d}s^2 = -\,\mathrm{d}X_1^2 - \mathrm{d}X_2^2 - \mathrm{d}X_3^2 - \mathrm{d}X_4^2 \tag{4.1}$$

then has a value which is independent of the orientation of the local system of co-ordinates, and is ascertainable by measurements of space and time. The magnitude of the linear element pertaining to points of the four-dimensional continuum in infinite proximity, we call $\mathrm{d}s$. If the $\mathrm{d}s$ belonging to the element $\mathrm{d}X_1$, $\mathrm{d}X_2$, $\mathrm{d}X_3$, and $\mathrm{d}X_4$ is positive, we follow Minkowski in calling it time-like; if it is negative, we call it space-like.
To the ”linear element” in question, or to the two infinitely proximate point-events, there will also correspond definite differentials $\mathrm{d}x_1$, $\mathrm{d}x_2$, $\mathrm{d}x_3$, and $\mathrm{d}x_4$ of the four-dimensional co-ordinates of any chosen system of reference. If this system, as well as the ”local” system, is given for the region under consideration, the $\mathrm{d}X_\nu$, will allow themselves to be represented here by definite linear homogeneous expressions

[1] The unit of time is to be chosen so that the velocity of light in vacuum as measured in the ”local” system of co-ordinates is to be equal to unity.

of the $\mathrm{d}x_\sigma$:–

$$\mathrm{d}X_\nu = \sum_\sigma \alpha_{\nu\sigma}\,\mathrm{d}x_\sigma\,. \tag{4.2}$$

Inserting these expressions in Equation 4.1, we obtain

$$\mathrm{d}s^2 = \sum_{\sigma\tau} g_{\sigma\tau}\,\mathrm{d}x_\sigma\,\mathrm{d}x_\tau\,, \tag{4.3}$$

where the $g_{\sigma\tau}$ will be functions of the x_σ. These can no longer be dependent on the orientation and the state of motion of the "local" system of co-ordinates, for $\mathrm{d}s^2$ is a quantity ascertainable by rod-clock measurement of point-events infinitely proximate in space-time, and defined independently of any particular choice of co-ordinates. The $g_{\sigma\tau}$ are to be chosen here so that $g_{\sigma\tau} = g_{\tau\sigma}$; the summation is to extend over all values of σ and τ, so that the sum consists of 4×4 terms, of which twelve are equal in pairs.

The case of the ordinary theory of relativity arises out of the case here considered, if it is possible, by reason of the particular relations of the $g_{\sigma\tau}$ in a finite region, to choose the system of reference in the finite region in such a way that the $g_{\sigma\tau}$ assume the constant values

$$\left.\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}\qquad\right\} \tag{4.4}$$

We shall find hereafter that the choice of such co-ordinates is, in general, not possible for a finite region.

From the considerations of chapter 2 and chapter 3 it follows that the quantities $g_{\sigma\tau}$ are to be regarded from the physical standpoint as the quantities which describe the gravitational field in relation to the chosen system of reference. For, if we now assume the special theory of relativity to apply to a certain four-dimensional region with the co-ordinates properly chosen, then the $g_{\sigma\tau}$ have the values given in Equation 4.4. A free material point then moves, relatively to this system, with uniform motion in a straight line. Then if we introduce new space-time co-ordinates x_1, x_2, x_3, x_4, by means of any substitution we choose, the $g^{\sigma\tau}$ in this new system will no longer be constants, but functions of space and time. At the same time the motion of the free material point will present itself in the new co-ordinates as a curvilinear non-uniform motion, and the law of this motion will be independent of the nature of the moving particle. We shall therefore interpret this motion as a motion under the influence of a gravitational field. We thus find the occurrence of a gravitational field connected with a space-time variability of the $g_{\sigma\tau}$. So, too, in the general case, when we are no longer able by a suitable choice of co-ordinates to apply the special theory of relativity to a finite region, we shall hold fast to the view that the $g_{\sigma\tau}$ describe the gravitational field.

Thus, according to the general theory of relativity, gravitation occupies an exceptional position with regard to other forces, particularly the electromagnetic forces, since the ten functions representing the gravitational field at the same time define the metrical properties of the space measured.

Part II.

MATHEMATICAL AIDS TO THE FORMULATION OF GENERALLY COVARIANT EQUATIONS

5. Introduction

Having seen in the foregoing that the general postulate of relativity leads to the requirement that the equations of physics shall be covariant in the face of any substitution of the co-ordinates x_1, x_2, x_3, x_4, we have to consider how such generally covariant equations can be found. We now turn to this purely mathematical task, and we shall find that in its solution a fundamental role is played by the invariant ds given in Equation 4.3, which, borrowing from GAUSS's theory of surfaces, we have called the "line element".
The fundamental idea of this general theory of covariants is the following:– Let certain things ("tensors") be defined with respect to any system of co-ordinates by a number of functions of the co-ordinates, called the "components" of the tensor. There are then certain rules by which these components can be calculated for a new system of co-ordinates, if they are known for the original system of co-ordinates, and if the transformation connecting the two systems is known. The things hereafter called tensors are further characterized by the fact that the equations of transformation for their components are linear and homogeneous. Accordingly, all the components in the new system vanish, if they all vanish in the original system. If, therefore, a law of nature is expressed by equating all the components of a tensor to zero, it is generally covariant. By examining the laws of the formation of tensors, we acquire the means of formulating generally covariant laws.

6. Contravariant and Covariant Four-vectors

Contravariant Four-vectors.– The linear element is defined by the four "components" $\mathrm{d}x_\nu$, for which the law of transformation is expressed by the equation

$$\mathrm{d}x'_\sigma = \sum_\nu \frac{\partial x'_\sigma}{\partial x_\nu}\,\mathrm{d}x_\nu \tag{6.1}$$

The $\mathrm{d}x'_\sigma$ are expressed as linear and homogeneous functions of the $\mathrm{d}x_\nu$. Hence we may look upon these co-ordinate differentials as the components of a "tensor" of the particular kind which we call a contravariant four-vector. Anything which is defined relatively to the system of co-ordinates by four quantities A^ν, and which is transformed by the same law

$$A'^\sigma = \sum_\nu \frac{\partial x'_\sigma}{\partial x_\nu} A^\nu\,, \tag{6.2}$$

we also call a contravariant four-vector. From Equation 6.2 it follows at once that the sums $A^\sigma \pm B^\sigma$ are also components of a four-vector, if A^σ and B^σ are such. Corresponding relations hold for all "tensors" subsequently to be introduced. (Rule for the addition and subtraction of tensors.)

Covariant Four-vectors.– We call four quantities A_ν the components of a covariant four-vector, if for any arbitrary choice of the contravariant four-vector B^ν

$$\sum_\nu A_\nu B^\nu = \text{Invariant.} \tag{6.3}$$

The law of transformation of a covariant four-vector follows from this definition. For if we replace B^ν on the right-hand side of the equation

$$\sum_\sigma A'_\sigma B'^\sigma = \sum_\nu A_\nu B^\nu$$

by the expression resulting from the inversion of Equation 6.2,

$$\sum_\sigma \frac{\partial x_\nu}{\partial x'_\sigma} B'^\sigma\,,$$

we obtain

$$\sum_\sigma B'^\sigma \sum_\nu \frac{\partial x_\nu}{\partial x'_\sigma} A_\nu = \sum_\sigma B'^\sigma A'_\sigma \,.$$

Since this equation is true for arbitrary values of the B'^σ, it follows that the law of transformation is

$$A'_\sigma = \sum_\nu \frac{\partial x_\nu}{\partial x'_\sigma} A_\nu \,. \tag{6.4}$$

Note on a Simplified Way of Writing the Expressions.–
A glance at the equations of this paragraph shows that there is always a summation with respect to the indices which occur twice under a sign of summation (e.g. the index ν in Equation 6.1), and *only* with respect to indices which occur twice. It is therefore possible, without loss of clearness, to omit the sign of summation. In its place we introduce the convention:–
If an index occurs twice in one term of an expression, it is always to be summed unless the contrary is expressly stated.
The difference between covariant and contravariant four-vectors lies in the law of transformation (Equation 6.4 or Equation 6.1 respectively). Both forms are tensors in the sense of the general remark above. Therein lies their importance. Following Ricci and Levi-Civita, we denote the contravariant character by placing the index above, the covariant by placing it below.

7. Tensors of the Second and Higher Ranks

Contravariant Tensors.– If we form all the sixteen products $A^{\mu\nu}$ of the components A^{μ} and B^{ν} of two contravaniant four-vectors

$$A^{\mu\nu} = A^{\mu} B^{\nu} \,, \tag{7.1}$$

then by Equation 7.1 and Equation 6.2 $A^{\mu\nu}$ satisfies the law of transformation

$$A'^{\sigma\tau} = \frac{\partial x'_{\sigma}}{\partial x_{\mu}} \frac{\partial x'_{\tau}}{\partial x_{\nu}} A^{\mu\nu} \,. \tag{7.2}$$

We call a thing which is described relatively to any system of reference by sixteen quantities, satisfying the law of transformation according to Equation 7.2, a contravariant tensor of the second rank. Not every such tensor allows itself to be formed in accordance with Equation 7.1 from two four-vectors, but it is easily shown that any given sixteen $A^{\mu\nu}$ can be represented as the sums of the $A^{\mu}B^{\nu}$ of four appropriately selected pairs of four-vectors. Hence we can prove nearly all the laws which apply to the tensor of the second rank defined by Equation 7.2 in the simplest manner by demonstrating them for the special tensors of the type Equation 7.1.

Contravariant Tensors of Any Rank.– It is clear that, on the lines of Equation 7.1 and Equation 7.2, contravariant tensors of the third and higher ranks may also be defined with 4^3 components, and so on. In the same way it follows from Equation 7.1 and Equation 7.2 that the contravariant four-vector may be taken in this sense as a contravariant tensor of the first rank.

Covariant Tensors.– On the other hand, if we take the sixteen products $A_{\mu\nu}$, of two covariant four-vectors A_{μ} and B_{ν},

$$A_{\mu\nu} = A_{\mu} B_{\nu} \,, \tag{7.3}$$

the law of transformation for these is

$$A'_{\sigma\tau} = \frac{\partial x_{\mu}}{\partial x'_{\sigma}} \frac{\partial x_{\nu}}{\partial x'_{\tau}} A_{\mu\nu} \,. \tag{7.4}$$

This law of transformation defines the covariant tensor of the second rank. All our previous remarks on contravariant tensors apply equally to covariant tensors.
Note.– It is convenient to treat the scalar (or invariant) both as a contravariant and a covariant tensor of zero rank.
Mixed Tensors.– We may also define a tensor of the second rank of the type

$$A_{\mu}{}^{\nu} = A_{\mu} B^{\nu} \tag{7.5}$$

which is covariant with respect to the index μ, and contravariant with respect to the index ν. Its law of transformation is

$$A'_{\sigma}{}^{\tau} = \frac{\partial x'_{\tau}}{\partial x_{\nu}} \frac{\partial x_{\mu}}{\partial x'_{\sigma}} A_{\mu}^{\nu} \,. \tag{7.6}$$

Naturally there are mixed tensors with any number of indices of covariant character, and any number of indices of contravariant character. Covariant and contravariant tensors may be looked upon as special cases of mixed tensors.
Symmetrical Tensors.– A contravariant, or a covariant tensor, of the second or higher rank is said to be symmetrical if two components, which are obtained the one from the other by the interchange of two indices, are equal. The tensor $A^{\mu\nu}$, or the tensor $A_{\mu\nu}$, is thus symmetrical if for any combination of the indices μ, ν,

$$A^{\mu\nu} = A^{\nu\mu} \,, \tag{7.7}$$

or respectively,

$$A_{\mu\nu} = A_{\nu\mu} \,. \tag{7.8}$$

It has to be proved that the symmetry thus defined is a property which is independent of the system of reference. It follows in fact from Equation 7.2, when Equation 7.7 is taken into consideration, that

$$A'^{\sigma\tau} = \frac{\partial x'_{\sigma}}{\partial x_{\mu}} \frac{\partial x'_{\tau}}{\partial x_{\nu}} A^{\mu\nu} = \frac{\partial x'_{\sigma}}{\partial x_{\mu}} \frac{\partial x'_{\tau}}{\partial x_{\nu}} A^{\nu\mu} = \frac{\partial x'_{\tau}}{\partial x_{\mu}} \frac{\partial x'_{\sigma}}{\partial x_{\nu}} A^{\mu\nu} = A'^{\tau\sigma} \,.$$

The last equation but one depends upon the interchange of the summation indices μ and ν, i.e. merely on a change of notation.
Antisymmetrical Tensors.– A contravariant or a covariant tensor of the second, third, or fourth rank is said to be antisymmetrical if two components, which are obtained the one from the other by the interchange of two indices, are equal and of opposite sign. The tensor $A^{\mu\nu}$, or the tensor $A_{\mu\nu}$, is therefore antisymmetrical, if always

$$A^{\mu\nu} = -A^{\nu\mu} \,, \tag{7.9}$$

or respectively,

$$A_{\mu\nu} = -A_{\nu\mu} \tag{7.10}$$

Of the sixteen components $A^{\mu\nu}$, the four components $A^{\mu\mu}$ vanish; the rest are equal and of opposite sign in pairs, so that there are only six components numerically different (a six-vector). Similarly we see that the antisymmetrical tensor of the third rank $A^{\mu\nu\sigma}$ has only four numerically different components, while the antisymmetrical tensor $A^{\mu\nu\sigma\tau}$ has only one. There are no antisymmetrical tensors of higher rank than the fourth in a continuum of four dimensions.

8. Multiplication of Tensors

Outer Multiplication of Tensors.– We obtain from the components of a tensor of rank n and of a tensor of rank m the components of a tensor of rank $n+m$ by multiplying each component of the one tensor by each component of the other. Thus, for example, the tensors T arise out of the tensors A and B of different kinds

$$\begin{aligned} T_{\mu\nu\sigma} &= A_{\mu\nu}B_{\sigma} , \\ T^{\mu\nu\sigma\tau} &= A^{\mu\nu}B^{\sigma\tau} , \\ T_{\mu\nu}^{\sigma\tau} &= A_{\mu\nu}B^{\sigma\tau} . \end{aligned}$$

The proof of the tensor character of T is given directly by the representations of Equation 7.1, Equation 7.3, Equation 7.5, or by the laws of transformation Equation 7.2, Equation 7.4, Equation 7.6. The equations Equation 7.1, Equation 7.3, Equation 7.5 are themselves examples of outer multiplication of tensors of the first rank.

"Contraction" of a Mixed Tensor.– From any mixed tensor we may form a tensor whose rank is less by two, by equating an index of covariant with one of contravariant character, and summing with respect to this index ("contraction"). Thus, for example, from the mixed tensor of the fourth rank $A_{\mu\nu}^{\sigma\tau}$, we obtain the mixed tensor of the second rank

$$A_{\nu}^{\tau} = A_{\mu\nu}^{\mu\tau} \left(= \sum_{\mu} A_{\mu\nu}^{\mu\tau} \right) ,$$

and from this, by a second contraction, the tensor of zero rank,

$$A = A_{\nu}^{\nu} = A_{\mu\nu}^{\mu\nu} .$$

The proof that the result of contraction really possesses the tensor character is given either by the representation of a tensor according to the generalization of Equation 7.5 in combination with Equation 6.3, or by the generalization of Equation 7.6.

Inner and Mixed Multiplication of Tensors.– These consist in a combination of outer multiplication with contraction.

Examples.– From the covariant tensor of the second rank $A_{\mu\nu}$ and the contravariant tensor of the first rank B^{σ} we form by outer multiplication the mixed tensor

$$D_{\mu\nu}^{\sigma} = A_{\mu\nu}B^{\sigma} .$$

On contraction with respect to the indices ν and σ, we obtain the covariant four-vector

$$D_\mu = D^\nu_{\mu\nu} = A_{\mu\nu} B^\nu \,.$$

This we call the inner product of the tensors $A_{\mu\nu}$ and B^σ. Analogously we form from the tensors $A_{\mu\nu}$, and $B^{\sigma\tau}$, by outer multiplication and double contraction, the inner product $A_{\mu\nu}B^{\mu\nu}$. By outer multiplication and one contraction, we obtain from $A_{\mu\nu}$ and $B^{\sigma\tau}$ the mixed tensor of the second rank $D^\tau_\mu = A_{\mu\nu}B^{\nu\tau}$. This operation may be aptly characterized as a mixed one, being "outer" with respect to the indices μ and τ, and "inner" with respect to the indices ν and σ.
We now prove a proposition which is often useful as evidence of tensor character. From what has just been explained, $A_{\mu\nu}B^{\mu\nu}$ is a scalar if $A_{\mu\nu}$, and $B^{\sigma\tau}$ are tensors. But we may also make the following assertion: If $A_{\mu\nu}B^{\mu\nu}$ is a scalar for *any choice of the tensor* $B^{\mu\nu}$, then $A_{\mu\nu}$ has tensor character. For, by hypothesis, for any substitution,

$$A'_{\sigma\tau} B'^{\sigma\tau} = A_{\mu\nu} B^{\mu\nu} \,.$$

But by an inversion of Equation 7.2

$$B^{\mu\nu} = \frac{\partial x_\mu}{\partial x'_\sigma} \frac{\partial x_\nu}{\partial x'_\tau} B'^{\sigma\tau} \,.$$

This, inserted in the above equation, gives

$$\left(A'_{\sigma\tau} - \frac{\partial x_\mu}{\partial x'_\sigma} \frac{\partial x_\nu}{\partial x'_\tau} A_{\mu\nu} \right) B'^{\sigma\tau} = 0 \,.$$

This can only be satisfied for arbitrary values of $B'^{\sigma\tau}$ if the bracket vanishes. The result then follows by Equation 7.4. This rule applies correspondingly to tensors of any rank and character, and the proof is analogous in all cases.
The rule may also be demonstrated in this form: If B^μ and C^ν are any vectors, and if, for all values of these, the inner product $A_{\mu\nu}B^\mu C^\nu$ is a scalar, then $A_{\mu\nu}$ is a covariant tensor. This latter proposition also holds good even if only the more special assertion is correct, that with any choice of the four-vector B^μ the inner product $A_{\mu\nu}B^\mu B^\nu$ is a scalar, if in addition it is known that $A_{\mu\nu}$ satisfies the condition of symmetry $A_{\mu\nu} = A_{\nu\mu}$. For by the method given above we prove the tensor character of $(A_{\mu\nu} + A_{\nu\mu})$, and from this the tensor character of $A_{\mu\nu}$ follows on account of symmetry. This also can be easily generalized to the case of covariant and contravariant tensors of any rank.
Finally, there follows from what has been proved, this law, which may also be generalized for any tensors: If for any choice of the four-vector B^ν the quantities $A_{\mu\nu}B^\nu$ form a tensor of the first rank, then $A_{\mu\nu}$ is a tensor of the second rank. For, if C^μ is any four-vector, then on account of the tensor character of $A_{\mu\nu}B^\nu$, the inner product $A_{\mu\nu}B^\nu C^\mu$ is a scalar for any choice of the two four-vectors B^ν and C^μ. From which the proposition follows.

9. Some Aspects of the Fundamental Tensor $g_{\mu\nu}$

The Covariant Fundamental Tensor.– In the invariant expression for the square of the linear element,

$$\mathrm{d}s^2 = g_{\mu\nu}\,\mathrm{d}x_\mu\,\mathrm{d}x_\nu\,,$$

the part played by the $\mathrm{d}x_\mu$ is that of a contravariant vector which may be chosen at will. Since further, $g_{\mu\nu} = g_{\nu\mu}$, it follows from the considerations of the preceding paragraph that $g_{\mu\nu}$ is a covariant tensor of the second rank. We call it the "fundamental tensor". In what follows we deduce some properties of this tensor which, it is true, apply to any tensor of the second rank. But as the fundamental tensor plays a special part in our theory, which has its physical basis in the peculiar effects of gravitation, it so happens that the relations to be developed are of importance to us only in the case of the fundamental tensor.

The Contravariant Fundamental Tensor.– If in the determinant formed by the elements $g_{\mu\nu}$, we take the co-factor of each of the $g_{\mu\nu}$ and divide it by the determinant $g = |g_{\mu\nu}|$, we obtain certain quantities $g^{\mu\nu} (= g^{\nu\mu})$ which, as we shall demonstrate, form a contravariant tensor.

By a known property of determinants

$$g_{\mu\sigma}g^{\nu\sigma} = \delta_\mu^{\ \nu}\,, \tag{9.1}$$

where the symbol $\delta_\mu^{\ \nu}$ denotes 1 or 0, according as $\mu = \nu$ or $\mu \neq \nu$.

Instead of the above expression for $\mathrm{d}s^2$ we may thus write

$$g_{\mu\sigma}\delta_\nu^{\ \sigma}\,\mathrm{d}x_\mu\,\mathrm{d}x_\nu\,,$$

or, by Equation 9.1

$$g_{\mu\sigma}g_{\nu\tau}g^{\sigma\tau}\,\mathrm{d}x_\mu\,\mathrm{d}x_\nu\,.$$

But, by the multiplication rules of the preceding paragraphs, the quantities

$$\mathrm{d}\xi_\sigma = g_{\mu\sigma}\,\mathrm{d}x_\mu$$

form a covariant four-vector, and in fact an arbitrary choosable four-vector, since the $\mathrm{d}x_\mu$ are arbitrary. By introducing this into our expression we obtain

$$\mathrm{d}s^2 = g^{\sigma\tau}\,\mathrm{d}\xi_\sigma\,\mathrm{d}\xi_\tau\,.$$

Since this, with the arbitrary choice of the vector $\mathrm{d}\xi_\sigma$, is a scalar, and $g^{\sigma\tau}$ by its definition is symmetrical in the indices σ and τ, it follows from the results of the preceding paragraph that $g^{\sigma\tau}$ is a contravariant tensor.
It further follows from Equation 9.1 that $\delta_\mu{}^\nu$ is also a tensor, which we may call the mixed fundamental tensor.
The Determinant of the Fundamental Tensor.– By the rule for the multiplication of determinants

$$|g_{\mu\alpha}g^{\alpha\nu}| = |g_{\mu\alpha}| \cdot |g^{\alpha\nu}| \,.$$

On the other hand

$$|g_{\mu\alpha}g^{\alpha\nu}| = \left|\delta_\mu{}^\nu\right| = 1 \,.$$

It therefore follows that

$$|g_{\mu\nu}| \cdot |g^{\mu\nu}| = 1 \,. \tag{9.2}$$

The Volume Scalar.– We seek first the law of transformation of the determinant $g = |g_{\mu\nu}|$. In accordance with Equation 7.4

$$g' = \left|\frac{\partial x_\mu}{\partial x'_\sigma}\frac{\partial x_\nu}{\partial x'_\tau} g_{\mu\nu}\right| \,.$$

Hence, by a double application of the rule for the multiplication of determinants, it follows that

$$g' = \left|\frac{\partial x_\mu}{\partial x'_\sigma}\right| \cdot \left|\frac{\partial x_\nu}{\partial x'_\tau}\right| \cdot |g_{\mu\nu}| = \left|\frac{\partial x_\mu}{\partial x'_\sigma}\right|^2 g \,,$$

or

$$\sqrt{g'} = \left|\frac{\partial x_\mu}{\partial x'_\sigma}\right| \cdot \sqrt{g} \,.$$

On the other hand, the law of transformation of the element of volume

$$\mathrm{d}\tau' = \int \mathrm{d}x_1 \,\mathrm{d}x_2 \,\mathrm{d}x_3 \,\mathrm{d}x_4$$

is, in accordance with the theorem of Jacobi,

$$\mathrm{d}\tau' = \left|\frac{\partial x'_\sigma}{\partial x_\mu}\right| \mathrm{d}\tau \,.$$

By multiplication of the last two equations, we obtain

$$\sqrt{g'} \cdot \mathrm{d}\tau' = \sqrt{g} \cdot \mathrm{d}\tau \,. \tag{9.3}$$

Instead of $\sqrt{g}$, we introduce in what follows the quantity $\sqrt{-g}$, Which is always real on account of the hyperbolic character of the space-time continuum. The invariant $\sqrt{-g}\,\mathrm{d}\tau$ is equal to the magnitude of the four-dimensional element of volume in the "local" system of reference, as measured with rigid rods and clocks

in the sense of the special theory of relativity.
Note on the Character of the Space-time Continuum.– Our assumption that the special theory of relativity can always be applied to an infinitely small region, implies that $\mathrm{d}s^2$ can always be expressed in accordance with Equation 4.1 by means of real quantities $\mathrm{d}X_1$, $\mathrm{d}X_2$, $\mathrm{d}X_3$, and $\mathrm{d}X_4$. If we denote by $\mathrm{d}\tau_0$ the "natural" element of volume $\mathrm{d}X_1$, $\mathrm{d}X_2$, $\mathrm{d}X_3$, $\mathrm{d}X_4$ then

$$\mathrm{d}\tau_0 = \sqrt{-g} \cdot \mathrm{d}\tau \,. \tag{9.4}$$

If $\sqrt{-g}$ were to vanish at a point of the four-dimensional continuum, it would mean that at this point an infinitely small "natural" volume would correspond to a finite volume in the co-ordinates. Let us assume that this is never the case. Then g cannot change sign. We will assume that, in the sense of the special theory of relativity, g always has a finite negative value. This is a hypothesis as to the physical nature of the continuum under consideration, and at the same time a convention as to the choice of co-ordinates.
But if $-g$ is always finite and positive, it is natural to settle the choice of co-ordinates *a posteriori* in such a way that this quantity is always equal to unity. We shall see later that by such a restriction of the choice of co-ordinates it is possible to achieve an important simplification of the laws of nature.
In place of Equation 9.3, we then have simply

$$\mathrm{d}\tau' = \mathrm{d}\tau \,,$$

from which, in view of JACOBI's theorem, it follows that

$$\left| \frac{\partial x'_\sigma}{\partial x_\mu} \right| = 1 \,. \tag{9.5}$$

Thus, with this choice of co-ordinates, only substitutions for which the determinant is unity are permissible.
But it would be erroneous to believe that this step indicates a partial abandonment of the general postulate of relativity. We do not ask "What are the laws of nature which are covariant in face of all substitutions for which the determinant is unity?", but our question is "What are the generally covariant laws of nature?". It is not until we have formulated these that we simplify their expression by a particular choice of the system of reference.
The Formation of New Tensors by Means of the Fundamental Tensor.– Inner, outer, and mixed multiplication of a tensor by the fundamental tensor give tensors of

different character and rank. For example

$$A^{\mu} = g^{\mu\sigma} A_{\sigma} \,,$$
$$A = g_{\mu\nu} A^{\mu\nu} \,.$$

The following forms may be specially noted:–

$$A^{\mu\nu} = g^{\mu\alpha} g^{\nu\beta} A_{\alpha\beta} \,,$$
$$A_{\mu\nu} = g_{\mu\alpha} g_{\nu\beta} A^{\alpha\beta}$$

(the "complements" of covariant and contravariant tensors respectively), and

$$B_{\mu\nu} = g_{\mu\nu} g^{\alpha\beta} A_{\alpha\beta} \,.$$

We call $B_{\mu\nu}$ the reduced tensor associated with $A_{\mu\nu}$.
Similarly,

$$B^{\mu\nu} = g^{\mu\nu} g_{\alpha\beta} A^{\alpha\beta} \,.$$

It may be noted that $g^{\mu\nu}$ is nothing more than the complement of $g_{\mu\nu}$, since

$$g^{\mu\alpha} g^{\nu\beta} g_{\alpha\beta} = g^{\mu\alpha} \delta_{\alpha}^{\ \nu} = g^{\mu\nu} \,.$$

10. The Equation of the Geodetic Line. The Motion of a Particle

As the linear element ds is defined independently of the system of co-ordinates, the line drawn between two points P and P' of the four-dimensional continuum in such a way that $\int \mathrm{d}s$ is stationary – a geodetic line – has a meaning which also is independent of the choice of co-ordinates. Its equation is

$$\delta \left\{ \int_P^{P'} \mathrm{d}s \right\} = 0\,. \tag{10.1}$$

Carrying out the variation in the usual way, we obtain from this equation four differential equations which define the geodetic line; this operation will be inserted here for the sake of completeness. Let λ be a function of the co-ordinates x_ν, and let this define a family of surfaces which intersect the required geodetic line as well as all the lines in immediate proximity to it which are drawn through the points P and P'. Any such line may then be supposed to be given by expressing its co-ordinates x_ν as functions of λ. Let the symbol δ indicate the transition from a point of the required geodetic to the point corresponding to the same λ on a neighbouring line. Then for Equation 10.1 we may substitute

$$\left.\begin{aligned} &\int_{\lambda_1}^{\lambda_2} \delta w \,\mathrm{d}\lambda = 0 \\ &\qquad w^2 = g_{\mu\nu} \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x_\nu}{\mathrm{d}\lambda}\,. \end{aligned}\right\} \tag{10.2}$$

But since

$$\delta w = \frac{1}{w} \cdot \left\{ \frac{1}{2} \cdot \frac{\partial g_{\mu\nu}}{\partial x_\sigma} \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x_\nu}{\mathrm{d}\lambda} \delta x_\sigma + g_{\mu\nu} \frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda} \delta \left(\frac{\mathrm{d}x_\nu}{\mathrm{d}\lambda} \right) \right\}\,,$$

and recognizing that

$$\delta \left(\frac{\mathrm{d}x_\nu}{\mathrm{d}\lambda} \right) = \frac{\mathrm{d}}{\mathrm{d}\lambda} \left(\delta x_\nu \right)\,,$$

we obtain from Equation 10.2, after a partial integration

$$\left.\begin{aligned} &\int_{\lambda_1}^{\lambda_2} \kappa_\sigma \delta x_\sigma \,\mathrm{d}\lambda = 0\,, \\ &\qquad\qquad \text{where} \\ &\kappa_\sigma = \frac{\mathrm{d}}{\mathrm{d}\lambda}\left\{\frac{g_{\mu\nu}}{w}\frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda}\right\} - \frac{1}{2w}\cdot\frac{\partial g_{\mu\nu}}{\partial x_\sigma}\frac{\mathrm{d}x_\mu}{\mathrm{d}\lambda}\frac{\mathrm{d}x_\nu}{\mathrm{d}\lambda}\,. \end{aligned}\right\} \tag{10.3}$$

Since the values of δx_σ are arbitrary, it follows from this that

$$\kappa_\sigma = 0 \tag{10.4}$$

are the equations of the geodetic line.
If $\mathrm{d}s$ does not vanish along the geodetic line we may choose the "length of the arc" s, measured along the geodetic line, for the parameter λ. Then $w = 1$, and in place of Equation 10.4 we obtain

$$g_{\mu\nu}\frac{\mathrm{d}^2 x_\mu}{\mathrm{d}s^2} + \frac{\partial g_{\mu\nu}}{\partial x_\sigma}\frac{\mathrm{d}x_\sigma}{\mathrm{d}s}\frac{\mathrm{d}x_\mu}{\mathrm{d}s} - \frac{1}{2}\cdot\frac{\partial g_{\mu\nu}}{\partial x_\sigma}\frac{\mathrm{d}x_\mu}{\mathrm{d}s}\frac{\mathrm{d}x_\nu}{\mathrm{d}s} = 0\,,$$

or, by a mere change of notation,

$$g_{\alpha\sigma}\frac{\mathrm{d}^2 x_\alpha}{\mathrm{d}s^2} + \begin{bmatrix}\mu\ \nu\\ \sigma\end{bmatrix}\frac{\mathrm{d}x_\mu}{\mathrm{d}s}\frac{\mathrm{d}x_\nu}{\mathrm{d}s} = 0\,, \tag{10.5}$$

where, following Christoffel, we have written

$$\begin{bmatrix}\mu\ \nu\\ \sigma\end{bmatrix} = \frac{1}{2}\cdot\left(\frac{\partial g_{\mu\sigma}}{\partial x_\nu} + \frac{\partial g_{\nu\sigma}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\sigma}\right)\,. \tag{10.6}$$

Finally, if we multiply Equation 10.5 by $g^{\sigma\tau}$ (outer multiplication with respect to τ, inner with respect to σ), we obtain the equations of the geodetic line in the form

$$\frac{\mathrm{d}^2 x_\tau}{\mathrm{d}s^2} + \begin{Bmatrix}\mu\ \nu\\ \tau\end{Bmatrix}\frac{\mathrm{d}x_\mu}{\mathrm{d}s}\frac{\mathrm{d}x_\nu}{\mathrm{d}s} = 0\,. \tag{10.7}$$

where, following Christoffel, we have set

$$\begin{Bmatrix}\mu\ \nu\\ \tau\end{Bmatrix} = g^{\tau\alpha}\begin{bmatrix}\mu\ \nu\\ \alpha\end{bmatrix}\,. \tag{10.8}$$

11. The Formation of Tensors by Differentiation

With the help of the equation of the geodetic line we can now easily deduce the laws by which new tensors can be formed from old by differentiation. By this means we are able for the first time to formulate generally covariant differential equations. We reach this goal by repeated application of the following simple law:– If in our continuum a curve is given, the points of which are specified by the arcual distance s measured from a fixed point on the curve, and if, further, ϕ is an invariant function of space, then $\frac{\mathrm{d}\phi}{\mathrm{d}s}$ is also an invariant.
The proof lies in this, that $\mathrm{d}s$ is an invariant as well as $\mathrm{d}\phi$.
As

$$\frac{\mathrm{d}\phi}{\mathrm{d}s} = \frac{\mathrm{d}\phi}{\mathrm{d}x_\mu}\frac{\mathrm{d}x_\mu}{\mathrm{d}s}\,,$$

therefore

$$\psi = \frac{\partial\phi}{\partial x_\mu}\frac{\mathrm{d}x_\mu}{\mathrm{d}s}$$

is also an invariant, and an invariant for all curves starting from a point of the continuum, that is, for any choice of the vector $\mathrm{d}x_\mu$. Hence it immediately follows that

$$A_\mu = \frac{\partial\phi}{\partial x_\mu} \tag{11.1}$$

is a covariant four-vector – the "gradient" of ϕ.
According to our rule, the differential quotient

$$\chi = \frac{\mathrm{d}\psi}{\mathrm{d}s}$$

taken on a curve, is similarly an invariant. Inserting the value of ψ, we obtain in the first place

$$\chi = \frac{\partial^2\phi}{\partial x_\mu\,\partial x_\nu}\frac{\mathrm{d}x_\mu}{\mathrm{d}s}\frac{\mathrm{d}x_\nu}{\mathrm{d}s} + \frac{\mathrm{d}\phi}{\mathrm{d}x_\mu}\frac{\mathrm{d}^2 x_\mu}{\mathrm{d}s^2}\,.$$

The existence of a tensor cannot be deduced from this forthwith. But if we may take the curve along which we have differentiated to be a geodetic, we obtain on substitution for $\frac{\mathrm{d}^2 x_\nu}{\mathrm{d}s^2}$ from Equation 10.7,

$$\chi = \left\{ \frac{\partial^2 \phi}{\partial x_\mu \, \partial x_\nu} - \begin{Bmatrix} \mu \ \nu \\ \tau \end{Bmatrix} \frac{\partial \phi}{\partial x_\tau} \right\} \frac{\mathrm{d}x_\mu}{\mathrm{d}s} \frac{\mathrm{d}x_\nu}{\mathrm{d}s} \,.$$

Since we may interchange the order of the differentiations, and since by Equation 10.8 and Equation 10.6 $\begin{Bmatrix} \mu \ \nu \\ \tau \end{Bmatrix}$ is symmetrical in μ and ν, it follows that the expression in brackets is symmetrical in μ and ν. Since a geodetic line can be drawn in any direction from a point of the continuum, and therefore $\frac{\mathrm{d}x_\mu}{\mathrm{d}s}$ is a four-vector with the ratio of its components arbitrary, it follows from the results of chapter 8 that

$$A_{\mu\nu} = \frac{\partial^2 \phi}{\partial x_\mu \, \partial x_\nu} - \begin{Bmatrix} \mu \ \nu \\ \tau \end{Bmatrix} \frac{\partial \phi}{\partial x_\tau} \tag{11.2}$$

is a covariant tensor of the second rank. We have therefore come to this result: from the covariant tensor of the first rank

$$A_\mu = \frac{\partial \phi}{\partial x_\mu}$$

we can, by differentiation, form a covariant tensor of the second rank

$$A_{\mu\nu} = \frac{\partial A_\mu}{\partial x_\nu} - \begin{Bmatrix} \mu \ \nu \\ \tau \end{Bmatrix} A_\tau \,. \tag{11.3}$$

We call the tensor $A_{\mu\nu}$ the "*extension*" (covariant derivative) of the tensor A_μ. In the first place we can readily show that the operation leads to a tensor, even if the vector A_μ cannot be represented as a gradient. To see this, we first observe that

$$\psi \cdot \frac{\partial \phi}{\partial x_\mu}$$

is a covariant vector, if ψ and ϕ are scalars. The sum of four such terms

$$S_\mu = \psi^{(1)} \frac{\partial \phi^{(1)}}{\partial x_\mu} + \psi^{(2)} \frac{\partial \phi^{(2)}}{\partial x_\mu} + \psi^{(3)} \frac{\partial \phi^{(3)}}{\partial x_\mu} + \psi^{(4)} \frac{\partial \phi^{(4)}}{\partial x_\mu} \,,$$

is also a covariant vector, if $\psi^{(1)}, \phi^{(1)}, \ldots, \psi^{(4)}, \phi^{(4)}$ are scalars. But it is clear that any covariant vector can be represented in the form S_μ. For, if A_μ is a vector whose components are any given functions of the x_ν, we have only to put (in terms of the selected system of co-ordinates)

$$\psi^{(1)} = A_1$$
$$\psi^{(2)} = A_2$$
$$\psi^{(3)} = A_3$$
$$\psi^{(4)} = A_4$$

and also

$$\phi^{(1)} = x_1$$
$$\phi^{(2)} = x_2$$
$$\phi^{(3)} = x_3$$
$$\phi^{(4)} = x_4\,,$$

in order to ensure that S_μ shall be equal to A_μ.
Therefore, in order to demonstrate that $A_{\mu\nu}$ is a tensor if *any* covariant vector is inserted on the right-hand side for A_μ, we only need show that this is so for the vector S_μ. But for this latter purpose it is sufficient, as a glance at the right-hand side of Equation 11.3 teaches us, to furnish the proof for the case

$$A_\mu = \psi \cdot \frac{\partial \phi}{\partial x_\mu}$$

Now the right-hand side of Equation 11.2 multiplied by ψ,

$$\psi \cdot \frac{\partial^2 \phi}{\partial x_\mu \, \partial x_\nu} - \begin{Bmatrix} \mu \; \nu \\ \tau \end{Bmatrix} \psi \cdot \frac{\partial \phi}{\partial x_\tau}$$

is a tensor. Similarly

$$\frac{\partial \psi}{\partial x_\mu} \frac{\partial \phi}{\partial x_\nu}$$

being the outer product of two vectors, is a tensor. By addition, there follows the tensor character of

$$\frac{\partial}{\partial x_\nu}\left(\psi \cdot \frac{\partial \phi}{\partial x_\mu}\right) - \begin{Bmatrix} \mu \; \nu \\ \tau \end{Bmatrix} \left(\psi \cdot \frac{\partial \phi}{\partial x_\tau}\right) .$$

As a glance at Equation 11.3 will show, this completes the demonstration for the vector

$$\psi \cdot \frac{\partial \phi}{\partial x_\mu}\,,$$

and consequently, from what has already been proved, for any vector A_μ.
By means of the extension of the vector, we may easily define the "extension" of a covariant tensor of any rank. This operation is a generalization of the extension of a vector. We restrict ourselves to the case of a tensor of the second rank, since this suffices to give a clear idea of the law of formation.
As has already been observed, any covariant tensor of the second rank can be represented[1] as the sum of tensors of the type $A_\mu B_\nu$. It will therefore be sufficient to deduce the expression for the extension of a tensor of this special type. By Equation 11.3 the expressions

$$\begin{gathered}
\frac{\partial A_\mu}{\partial x_\sigma} - \begin{Bmatrix} \sigma\ \mu \\ \tau \end{Bmatrix} A_\tau\,, \\
\frac{\partial B_\nu}{\partial x_\sigma} - \begin{Bmatrix} \sigma\ \nu \\ \tau \end{Bmatrix} B_\tau
\end{gathered}$$

are tensors. On outer multiplication of the first by B_ν, and of the second by A_μ, we obtain in each case a tensor of the third rank. By adding these, we have the tensor of the third rank

$$A_{\mu\nu\sigma} = \frac{\partial A_{\mu\nu}}{\partial x_\sigma} - \begin{Bmatrix} \sigma\ \mu \\ \tau \end{Bmatrix} A_{\tau\nu} - \begin{Bmatrix} \sigma\ \nu \\ \tau \end{Bmatrix} A_{\mu\tau} \tag{11.4}$$

where we have put $A_{\mu\nu} = A_\mu B_\nu$. As the right-hand side of Equation 11.4 is linear and homogeneous in the $A_{\mu\nu}$ and their first derivatives, this law of formation leads to a tensor, not only in the case of a tensor of the type $A_\mu B_\nu$, but also in the case of a sum of such tensors, i.e. in the case of any covariant tensor of the second rank. We call $A_{\mu\nu\sigma}$ the extension of the tensor $A_{\mu\nu}$.
It is clear that Equation 11.3 and Equation 11.1 concern only special cases of extension (the extension of the tensors of rank one and zero respectively).
In general, all special laws of formation of tensors are included in Equation 11.4 in combination with the multiplication of tensors.

[1] By outer multiplication of the vector with arbitrary components A_{11}, A_{12}, A_{13}, A_{14} by the vector with components 1, 0, 0, 0, we produce a tensor with components

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} & A_{14} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

By the addition of four tensors of this type, we obtain the tensor $A_{\mu\nu}$ with any assigned components.

12. Some Cases of Special Importance

The Fundamental Tensor.– We will first prove some lemmas which will be useful hereafter. By the rule for the differentiation of determinants

$$\mathrm{d}g = g^{\mu\nu} g \, \mathrm{d}g_{\mu\nu} = -g_{\mu\nu} g \, \mathrm{d}g^{\mu\nu} \,. \tag{12.1}$$

the last member is obtained from the last but one, if we bear in mind that $g_{\mu\nu} g^{\mu'\nu} = \delta_{\mu}{}^{\mu'}$, so that $g_{\mu\nu} g^{\mu\nu} = 4$, and consequently

$$g_{\mu\nu} \, \mathrm{d}g^{\mu\nu} + g^{\mu\nu} g \, \mathrm{d}g_{\mu\nu} = 0 \,.$$

From Equation 12.1, it follows that

$$\frac{1}{\sqrt{-g}} \frac{\partial \sqrt{-g}}{\partial x_\sigma} = \frac{1}{2} \cdot \frac{\partial \log(\sqrt{-g})}{\partial x_\sigma} = \frac{1}{2} \cdot g^{\mu\nu} \frac{\partial g_{\mu\nu}}{\partial x_\sigma} = -\frac{1}{2} \cdot g_{\mu\nu} \frac{\partial g^{\mu\nu}}{\partial x_\sigma} \,. \tag{12.2}$$

Further, from

$$g_{\mu\sigma} g^{\nu\sigma} = \delta_{\mu}{}^{\nu}$$

it follows on differentiation that

$$\left.\begin{aligned} g_{\mu\sigma} \, \mathrm{d}g^{\nu\sigma} &= -g^{\nu\sigma} \, \mathrm{d}g_{\mu\sigma} \\ \text{and} \qquad\qquad & \\ g_{\mu\sigma} \frac{\partial g^{\nu\sigma}}{\partial x_\lambda} &= -g^{\nu\sigma} \frac{\partial g_{\mu\sigma}}{\partial x_\lambda} \,. \end{aligned}\right\} \tag{12.3}$$

From these, by mixed multiplication by $g^{\sigma\tau}$ and $g_{\nu\lambda}$ respectively, and a change of notation for the indices, we have

$$\left.\begin{aligned} \mathrm{d}g^{\mu\nu} &= -g^{\mu\alpha} g^{\nu\beta} \, \mathrm{d}g_{\alpha\beta} \,, \\ \frac{\partial g^{\mu\nu}}{\partial x_\sigma} &= -g^{\mu\alpha} g^{\nu\beta} \frac{\partial g_{\alpha\beta}}{\partial x_\sigma} \end{aligned}\right\} \tag{12.4}$$

and

$$\left.\begin{aligned} \mathrm{d}g_{\mu\nu} &= -g_{\mu\alpha}g_{\nu\beta}\,\mathrm{d}g^{\alpha\beta} \\ \frac{\partial g_{\mu\nu}}{\partial x_\sigma} &= -g_{\mu\alpha}g_{\nu\beta}\frac{\partial g^{\alpha\beta}}{\partial x_\sigma}\,. \end{aligned}\right\} \tag{12.5}$$

The Equation 12.4 admits of a transformation, of which we also have frequently to make use: From Equation 10.6

$$\frac{\partial g_{\alpha\beta}}{\partial x_\sigma} = \begin{bmatrix}\alpha\ \sigma\\ \beta\end{bmatrix} + \begin{bmatrix}\beta\ \sigma\\ \alpha\end{bmatrix}\,. \tag{12.6}$$

Inserting this in the second formula of Equation 12.4, we obtain, in view of Equation 10.8

$$\frac{\partial g^{\mu\nu}}{\partial x_\sigma} = -\left(g^{\mu\tau}\begin{Bmatrix}\tau\ \sigma\\ \nu\end{Bmatrix} + g^{\nu\tau}\begin{Bmatrix}\tau\ \sigma\\ \mu\end{Bmatrix}\right)\,. \tag{12.7}$$

Substituting the right-hand side of Equation 12.7 in Equation 12.2, we have

$$\frac{1}{\sqrt{-g}}\cdot\frac{\partial\sqrt{-g}}{\partial x_\sigma} = \begin{Bmatrix}\mu\ \sigma\\ \mu\end{Bmatrix}\,. \tag{12.8}$$

The "Divergence" of a Contravariant Vector.– If we take the inner product of Equation 11.3 by the contravariant fundamental tensor $g^{\mu\nu}$, the right-hand side, after a transformation of the first term, assumes the form

$$\frac{\partial}{\partial x_\nu}\left(g^{\mu\nu}A_\mu\right) - A_\mu\frac{\partial g^{\mu\nu}}{\partial x_\nu} - \frac{1}{2}\cdot g^{\tau\alpha}\left(\frac{\partial g_{\mu\alpha}}{\partial x_\nu} + \frac{\partial g_{\nu\alpha}}{\partial x_\mu} - \frac{\partial g_{\mu\nu}}{\partial x_\alpha}\right)g^{\mu\nu}A_\tau\,.$$

In accordance with Equation 12.4 and Equation 12.2, the last term of this expression may be written as

$$\frac{1}{2}\cdot\frac{\partial g^{\tau\nu}}{\partial x_\nu}A_\tau + \frac{1}{2}\cdot\frac{\partial g^{\tau\mu}}{\partial x_\mu}A_\tau + \frac{1}{\sqrt{-g}}\cdot\frac{\partial\sqrt{-g}}{\partial x_\alpha}g^{\mu\nu}A_\tau\,.$$

As the symbols of the indices of summation are immaterial, the first two terms of this expression cancel the second of the one above. If we then write

$$g^{\mu\nu}A_\mu = A^\nu$$

so that A^ν like A_μ is an arbitrary vector, we finally obtain

$$\Phi = \frac{1}{\sqrt{-g}}\cdot\frac{\partial}{\partial x_\nu}\left(\sqrt{-g}\cdot A^\nu\right)\,. \tag{12.9}$$

This scalar is the *divergence* of the contravariant vector A^ν.
The "Curl" of a Covariant Vector.– The second term in Equation 11.3 is symmetrical in the indices μ and ν. Therefore $A_{\mu\nu} - A_{\nu\mu}$ is a particularly simply constructed antisymmetrical tensor. We obtain

$$B_{\mu\nu} = \frac{\partial A_\mu}{\partial x_\nu} - \frac{\partial A_\nu}{\partial x_\mu} \tag{12.10}$$

Antisymmetrical Extension of a Six-vector.– Applying Equation 11.4 to an antisymmetrical tensor of the second rank $A_{\mu\nu}$, forming in addition the two equations which arise through cyclic permutations of the indices, and adding these three equations, we obtain the tensor of the third rank

$$B_{\mu\nu\sigma} = A_{\mu\nu\sigma} + A_{\nu\sigma\mu} + A_{\sigma\mu\nu} = \frac{\partial A_{\mu\nu}}{\partial x_\sigma} + \frac{\partial A_{\nu\sigma}}{\partial x_\mu} + \frac{\partial A_{\sigma\mu}}{\partial x_\nu}\,, \tag{12.11}$$

which it is easy to prove is antisymmetrical.
The Divergence of a Six-vector.– Taking the mixed product of Equation 11.4 by $g^{\mu\alpha}g^{\nu\beta}$, we also obtain a tensor. The first term on the right-hand side of Equation 11.4 may be written in the form

$$\frac{\partial}{\partial x_\sigma}\left(g^{\mu\alpha}g^{\nu\beta}A_{\mu\nu}\right) - g^{\mu\alpha}\frac{\partial g^{\nu\beta}}{\partial x_\sigma}A_{\mu\nu} - g^{\nu\beta}\frac{\partial g^{\mu\alpha}}{\partial x_\sigma}A_{\mu\nu}$$

If we write $A^{\alpha\beta}_\sigma$ for $g^{\mu\alpha}g^{\nu\beta}A_{\mu\nu\sigma}$ and $A^{\alpha\beta}$ for $g^{\mu\alpha}g^{\nu\beta}A_{\mu\nu}$, and in the transformed first term replace

$$\frac{\partial g^{\nu\beta}}{\partial x_\sigma} \quad \text{und} \quad \frac{\partial g^{\mu\alpha}}{\partial x_\sigma}$$

by their values as given by Equation 12.7, there results from the right-hand side of Equation 11.4 an expression consisting of seven terms, of which four cancel, and there remains

$$A^{\alpha\beta}_\sigma = \frac{\partial A^{\alpha\beta}}{\partial x_\sigma} + \begin{Bmatrix}\sigma\ \gamma\\ \alpha\end{Bmatrix} A^{\gamma\beta} + \begin{Bmatrix}\sigma\ \gamma\\ \beta\end{Bmatrix} A^{\alpha\gamma}\,. \tag{12.12}$$

This is the expression for the extension of a contravariant tensor of the second rank, and corresponding expressions for the extension of contravariant tensors of higher and lower rank may also be formed.
We note that in an analogous way we may also form the extension of a mixed tensor:–

$$A_{\mu\sigma}^{\ \ \alpha} = \frac{\partial A_\mu^{\ \alpha}}{\partial x_\sigma} - \begin{Bmatrix}\sigma\ \mu\\ \tau\end{Bmatrix} A_\tau^{\ \alpha} + \begin{Bmatrix}\sigma\ \tau\\ \alpha\end{Bmatrix} A_\mu^{\ \tau}\,. \tag{12.13}$$

On contracting Equation 12.12 with respect to the indices β and σ (inner multiplication by $\delta_\beta{}^\sigma$), we obtain the vector

$$A^\alpha = \frac{\partial A^{\alpha\beta}}{\partial x_\beta} + \left\{ \begin{matrix} \beta\ \gamma \\ \beta \end{matrix} \right\} A^{\alpha\gamma} + \left\{ \begin{matrix} \beta\ \gamma \\ \alpha \end{matrix} \right\} A^{\gamma\beta} .$$

On account of the symmetry of $\left\{ \begin{matrix} \beta\ \gamma \\ \alpha \end{matrix} \right\}$ with respect to the indices β and γ, the third term on the right-hand side vanishes, if $A^{\alpha\beta}$ is, as we will assume, an antisymmetrical tensor. The second term allows itself to be transformed in accordance with Equation 12.8. Thus we obtain

$$A^\alpha = \frac{1}{\sqrt{-g}} \cdot \frac{\partial \left(\sqrt{-g} \cdot A^{\alpha\beta} \right)}{\partial x_\beta} . \tag{12.14}$$

This is the expression for the divergence of a contravariant six-vector.

The Divergence of a Mixed Tensor of the Second Rank.– Contracting Equation 12.13 with respect to the indices α and σ, and taking Equation 12.8 into consideration, we obtain

$$\sqrt{-g} \cdot A_\mu = \frac{\partial \left(\sqrt{-g} \cdot A_\mu{}^\sigma \right)}{\partial x_\sigma} - \left\{ \begin{matrix} \sigma\ \mu \\ \tau \end{matrix} \right\} \sqrt{-g} \cdot A_\tau{}^\sigma . \tag{12.15}$$

If we introduce the contravariant tensor $A^{\rho\sigma} = g^{\rho\tau} A_\tau{}^\sigma$ in the last term, it assumes the form

$$- \left[\begin{matrix} \sigma\ \mu \\ \rho \end{matrix} \right] \sqrt{-g} \cdot A^{\rho\sigma} .$$

If, further, the tensor $A^{\rho\sigma}$ is symmetrical, this reduces to

$$-\frac{1}{2} \cdot \sqrt{-g} \cdot \frac{\partial g_{\rho\sigma}}{\partial x_\mu} A^{\rho\sigma} .$$

Had we introduced, instead of $A^{\rho\sigma}$, the covariant tensor $A_{\rho\sigma} = g_{\rho\alpha} g_{\sigma\beta} A^{\alpha\beta}$, which is also symmetrical, the last term, by virtue of Equation 12.4, would assume the form

$$\frac{1}{2} \cdot \sqrt{-g} \cdot \frac{\partial g^{\rho\sigma}}{\partial x_\mu} A_{\rho\sigma} .$$

In the case of symmetry in question, Equation 12.15 may therefore be replaced by the two forms

$$\sqrt{-g} \cdot A_\mu = \frac{\partial \left(\sqrt{-g} \cdot A_\mu^\sigma \right)}{\partial x_\sigma} - \frac{1}{2} \cdot \frac{\partial g_{\rho\sigma}}{\partial x_\mu} \sqrt{-g} \cdot A^{\rho\sigma} \tag{12.16a}$$

and

$$\sqrt{-g} \cdot A_\mu = \frac{\partial \left(\sqrt{-g} \cdot A_\mu^\sigma \right)}{\partial x_\sigma} - \frac{1}{2} \cdot \frac{\partial g^{\rho\sigma}}{\partial x_\mu} \sqrt{-g} \cdot A_{\rho\sigma} \tag{12.16b}$$

which we have to employ later on.

13. The Riemann-Christoffel Tensor

We now seek the tensor which can be obtained from the fundamental tensor *alone*, by differentiation. At first sight the solution seems obvious. We place the fundamental tensor of the $g_{\mu\nu}$ in Equation 11.4 instead of any given tensor $A_{\mu\nu}$, and thus have a new tensor, namely, the extension of the fundamental tensor. But we easily convince ourselves that this extension vanishes identically. We reach our goal, however, in the following way. In Equation 11.4 place

$$A_{\mu\nu} = \frac{\partial A_\mu}{\partial x_\nu} - \begin{Bmatrix} \mu\ \nu \\ \rho \end{Bmatrix} A_\rho \,,$$

i.e., the extension of the four-vector A_μ. Then (with a somewhat different naming of the indices) we get the tensor of the third rank

$$\begin{aligned} A_{\mu\sigma\tau} = & \frac{\partial^2 A_\mu}{\partial x_\sigma \, \partial x_\tau} \\ & - \begin{Bmatrix} \mu\ \sigma \\ \rho \end{Bmatrix} \frac{\partial A_\rho}{\partial x_\tau} - \begin{Bmatrix} \mu\ \tau \\ \rho \end{Bmatrix} \frac{\partial A_\rho}{\partial x_\sigma} - \begin{Bmatrix} \sigma\ \tau \\ \rho \end{Bmatrix} \frac{\partial A_\mu}{\partial x_\rho} \\ & + \left[-\frac{\partial}{\partial x_\tau} \begin{Bmatrix} \mu\ \sigma \\ \rho \end{Bmatrix} + \begin{Bmatrix} \mu\ \tau \\ \alpha \end{Bmatrix} \begin{Bmatrix} \alpha\ \sigma \\ \rho \end{Bmatrix} + \begin{Bmatrix} \sigma\ \tau \\ \alpha \end{Bmatrix} \begin{Bmatrix} \alpha\ \mu \\ \rho \end{Bmatrix} \right] A_\rho \,. \end{aligned}$$

This expression suggests forming the tensor $A_{\mu\sigma\tau} - A_{\mu\tau\sigma}$. For, if we do so, the following terms of the expression for $A_{\mu\sigma\tau}$ cancel those of $A_{\mu\tau\sigma}$, the first, the fourth, and the member corresponding to the last term in square brackets; because all these are symmetrical in σ and τ. The same holds good for the sum of the second and third terms. Thus we obtain

$$A_{\mu\sigma\tau} - A_{\mu\tau\sigma} = B^\rho_{\mu\sigma\tau} A_\rho \,, \tag{13.1}$$

where

$$\left. \begin{aligned} B^\rho_{\mu\sigma\tau} = & -\frac{\partial}{\partial x_\tau} \begin{Bmatrix} \mu\ \sigma \\ \rho \end{Bmatrix} + \frac{\partial}{\partial x_\sigma} \begin{Bmatrix} \mu\ \tau \\ \rho \end{Bmatrix} \\ & - \begin{Bmatrix} \mu\ \sigma \\ \alpha \end{Bmatrix} \begin{Bmatrix} \alpha\ \tau \\ \rho \end{Bmatrix} + \begin{Bmatrix} \mu\ \tau \\ \alpha \end{Bmatrix} \begin{Bmatrix} \alpha\ \sigma \\ \rho \end{Bmatrix} \,. \end{aligned} \right\} \tag{13.2}$$

The essential feature of the result is that on the right side of Equation 13.1 the A_ρ occur alone, without their derivatives. From the tensor character of $A_{\mu\sigma\tau} - A_{\mu\tau\sigma}$

in conjunction with the fact that A_ρ is an arbitrary vector, it follows, by reason of chapter 8, that $B^\rho_{\mu\sigma\tau}$ is a tensor (the RIEMANN-CHRISTOFFEL tensor).
The mathematical importance of this tensor is as follows: If the continuum is of such a nature that there is a co-ordinate system with reference to which the $g_{\mu\nu}$ are constants, then all the $B^\rho_{\mu\sigma\tau}$ vanish. If we choose any new system of coordinates in place of the original ones, the $g_{\mu\nu}$ referred thereto will not be constants, but in consequence of its tensor nature, the transformed components of $B^\rho_{\mu\sigma\tau}$ will still vanish in the new system. Thus the vanishing of the RIEMANN tensor is a necessary condition that, by an appropriate choice of the system of reference, the $g_{\mu\nu}$ may be constants. In our problem this corresponds to the case in which,[1] with a suitable choice of the system of reference, the special theory of relativity holds good for a *finite* region of the continuum.
Contracting Equation 13.2 with respect to the indices τ and ρ we obtain the covariant tensor of second rank

$$\left.\begin{aligned} G_{\mu\nu} &= B^\rho_{\mu\sigma\rho} = R_{\mu\nu} + S_{\mu\nu} \\ &\quad \text{where} \\ R_{\mu\nu} &= -\frac{\partial}{\partial x_\alpha}\begin{Bmatrix}\mu\ \nu \\ \alpha\end{Bmatrix} + \begin{Bmatrix}\mu\ \alpha \\ \beta\end{Bmatrix}\begin{Bmatrix}\nu\ \beta \\ \alpha\end{Bmatrix} \\ S_{\mu\nu} &= \frac{\partial^2 (\log\sqrt{-g})}{\partial x_\mu\, \partial x_\nu} - \begin{Bmatrix}\mu\ \nu \\ \alpha\end{Bmatrix}\frac{\partial (\log\sqrt{-g})}{\partial x_\alpha}\,. \end{aligned}\right\} \tag{13.3}$$

Note on the Choice of Co-ordinates.– It has already been observed in chapter 9, in connexion with Equation 9.4, that the choice of co-ordinates may with advantage be made so that $\sqrt{-g} = 1$. A glance at the equations obtained in the last two sections shows that by such a choice the laws of formation of tensors undergo an important simplification. This applies particularly to $G_{\mu\nu}$, the tensor just developed, which plays a fundamental part in the theory to be set forth. For this specialization of the choice of co-ordinates brings about the vanishing of $S_{\mu\nu}$, so that the tensor $G_{\mu\nu}$ reduces to $R_{\mu\nu}$.
On this account I shall hereafter give all relations in the simplified form which this specialization of the choice of coordinates brings with it. It will then be an easy matter to revert to the *generally* covariant equations, if this seems desirable in a special case.

[1]The mathematicians have proved that this is also a *sufficient* condition.

Part III.

THEORY OF THE GRAVITATIONAL FIELD

14. Equations of Motion of a Material Point in the Gravitational Field. Expression for the Field-components of Gravitation

A freely movable body not subjected to external forces moves, according to the special theory of relativity, in a straight line and uniformly. This is also the case, according to the general theory of relativity, for a part of four-dimensional space in which the system of co-ordinates K_0, may be, and is, so chosen that they have the special constant values given in Equation 4.4.
If we consider precisely this movement from any chosen system of co-ordinates K_1, the body, observed from K_1, moves, according to the considerations in chapter 2, in a gravitational field. The law of motion with respect to K_1 results without difficulty from the following consideration. With respect to K_0 the law of motion corresponds to a four-dimensional straight line, i.e. to a geodetic line. Now since the geodetic line is defined independently of the system of reference, its equations will also be the equation of motion of the material point with respect to K_1. If we set

$$\Gamma^{\tau}_{\mu\nu} = -\begin{Bmatrix} \mu\ \nu \\ \tau \end{Bmatrix} , \tag{14.1}$$

the equation of the motion of the point with respect to K_1, becomes

$$\frac{\mathrm{d}^2 x_\tau}{\mathrm{d}s^2} = \Gamma^{\tau}_{\mu\nu} \frac{\mathrm{d}x_\mu}{\mathrm{d}s} \frac{\mathrm{d}x_\nu}{\mathrm{d}s} . \tag{14.2}$$

We now make the assumption, which readily suggests itself, that this covariant system of equations also defines the motion of the point in the gravitational field in the case when there is no system of reference K_0, with respect to which the special theory of relativity holds good in a finite region. We have all the more justification for this assumption as Equation 14.2 contains only *first* derivatives of the $g_{\mu\nu}$, between which even in the special case of the existence of K_0, no relations subsist.[1]

If the $\Gamma^{\tau}_{\mu\nu}$ vanish, then the point moves uniformly in a straight line. These quantities therefore condition the deviation of the motion from uniformity. They are the components of the gravitational field.

[1] It is only between the second (and first) derivatives that, by chapter 13, the relations $B^{\rho}_{\mu\sigma\tau} = 0$ subsist.

15. The Field Equations of Gravitation in the Absence of Matter

We make a distinction hereafter between "gravitational field" and "matter" in this way, that we denote everything but the gravitational field as "matter". Our use of the word therefore includes not only matter in the ordinary sense, but the electromagnetic field as well.

Our next task is to find the field equations of gravitation in the absence of matter. Here we again apply the method employed in the preceding paragraph in formulating the equations of motion of the material point. A special case in which the required equations must in any case be satisfied is that of the special theory of relativity, in which the $g_{\mu\nu}$ have certain constant values. Let this be the case in a certain finite space in relation to a definite system of co-ordinates K_0. Relatively to this system all the components of the RIEMANN tensor $B^{\rho}_{\mu\sigma\tau}$ defined in Equation 13.2, vanish. For the space under consideration they then vanish, also in any other system of co-ordinates.

Thus the required equations of the matter-free gravitational field must in any case be satisfied if all $B^{\rho}_{\mu\sigma\tau}$ vanish. But this condition goes too far. For it is clear that, e.g., the gravitational field generated by a material point in its environment certainly cannot be "transformed away" by any choice of the system of co-ordinates, i.e. it cannot be transformed to the case of constant $g_{\mu\nu}$.

This prompts us to require for the matter-free gravitational field that the symmetrical tensor $G_{\mu\nu}$, derived from the tensor $B^{\rho}_{\mu\sigma\tau}$ shall vanish. Thus we obtain ten equations for the ten quantities $g_{\mu\nu}$, which are satisfied in the special case of the vanishing of all $B^{\rho}_{\mu\sigma\tau}$. With the choice which we have made of a system of co-ordinates, and taking Equation 13.3 into consideration, the equations for the matter-free field are

$$\left.\begin{aligned}\frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x_{\alpha}} + \Gamma^{\alpha}_{\mu\beta} \cdot \Gamma^{\beta}_{\nu\alpha} = 0 \\ \sqrt{-g} = 1\,.\end{aligned}\right\} \tag{15.1}$$

It must be pointed out that there is only a minimum of arbitrariness in the choice of these equations. For besides $G_{\mu\nu}$ there is no tensor of second rank which is formed from the $g_{\mu\nu}$ and its derivatives, contains no derivations higher than second,

and is linear in these derivatives.[1]
These equations, which proceed, by the method of pure mathematics, from the requirement of the general theory of relativity, give us, in combination with the equations of motion Equation 14.2, to a first approximation NEWTON's law of attraction, and to a second approximation the explanation of the motion of the perihelion of the planet Mercury discovered by LEVERRIER (as it remains after corrections for perturbation have been made). These facts must, in my opinion, be taken as a convincing proof of the correctness of the theory.

[1]Properly speaking, this can be affirmed only of the tensor

$$G_{\mu\nu} + \lambda g_{\mu\nu} g^{\alpha\beta} G_{\alpha\beta} \,,$$

where λ a constant. If, however, we set this tensor $= 0$, we come back again to the equations $G_{\mu\nu} = 0$.

16. The Hamiltonian Function for the Gravitational Field. Laws of Momentum and Energy

To show that the field equations correspond to the laws of momentum and energy, it is most convenient to write them in the following Hamiltonian form:–

$$\left.\begin{aligned} \delta\left\{\int H \,\mathrm{d}\tau\right\} &= 0 \\ H &= g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} \\ \sqrt{-g} &= 1\,. \end{aligned}\right\} \tag{16.1}$$

where, on the boundary of the finite four-dimensional region of integration which we have in view, the variations vanish.
We first have to show that the form Equation 16.1 is equivalent to the equations according to Equation 15.1. For this purpose we regard H as a function of the $g^{\mu\nu}$ and the

$$g^{\mu\nu}_{\sigma} \quad \left(= \frac{\partial g^{\mu\nu}}{\partial x_\sigma}\right)\,.$$

Then in the first place

$$\begin{aligned} \delta H &= \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\delta g^{\mu\nu} + 2\cdot g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\delta\Gamma^{\beta}_{\nu\alpha} \\ &= -\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\delta g^{\mu\nu} + 2\cdot\Gamma^{\alpha}_{\mu\beta}\delta\left(g^{\mu\nu}\Gamma^{\beta}_{\nu\alpha}\right)\,. \end{aligned}$$

But

$$\delta\left(g^{\mu\nu}\Gamma^{\beta}_{\nu\alpha}\right) = -\frac{1}{2}\cdot\delta\left[g^{\mu\nu}g^{\beta\lambda}\cdot\left(\frac{\partial g_{\nu\lambda}}{\partial x_\alpha} + \frac{\partial g_{\alpha\lambda}}{\partial x_\nu} - \frac{\partial g_{\alpha\nu}}{\partial x_\lambda}\right)\right]\,.$$

The terms arising from the last two terms in round brackets are of different sign, and result from each other (since the denomination of the summation indices is immaterial) through interchange of the indices μ and β. They cancel each other in the expression for δH, because they are multiplied by the quantity $\Gamma^{\alpha}_{\mu\beta}$, which is symmetrical with respect to the indices μ and β. Thus there remains only the first term in round brackets to be considered, so that, taking Equation 12.4 into account, we obtain

$$\delta H = -\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha}\delta g^{\mu\nu} + \Gamma^{\alpha}_{\mu\beta}\delta g^{\mu\beta}_{\alpha}\,.$$

Thus

$$\left.\begin{aligned} \frac{\partial H}{\partial g^{\mu\nu}} &= -\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} \\ \frac{\partial H}{\partial g^{\mu\nu}_{\sigma}} &= \Gamma^{\sigma}_{\mu\nu}\,. \end{aligned}\right\} \tag{16.2}$$

Carrying out the variation in Equation 16.1, we get in the first place

$$\frac{\partial}{\partial x_\alpha}\left(\frac{\partial H}{\partial g^{\mu\nu}_{\alpha}}\right) - \frac{\partial H}{\partial g^{\mu\nu}} = 0\,, \tag{16.3}$$

which, on account of Equation 16.2, agrees with Equation 15.1, as was to be proved. If we multiply Equation 16.3 by $g^{\mu\nu}_{\sigma}$, then because

$$\frac{\partial g^{\mu\nu}_{\sigma}}{\partial x_\alpha} = \frac{\partial g^{\mu\nu}_{\alpha}}{\partial x_\sigma}$$

and, consequently,

$$g^{\mu\nu}_{\sigma}\frac{\partial}{\partial x_\alpha}\left(\frac{\partial H}{\partial g^{\mu\nu}_{\alpha}}\right) = \frac{\partial}{\partial x_\alpha}\left(g^{\mu\nu}_{\sigma}\frac{\partial H}{\partial g^{\mu\nu}_{\alpha}}\right) - \frac{\partial H}{\partial g^{\mu\nu}_{\alpha}}\frac{\partial g^{\mu\nu}_{\alpha}}{\partial x_\sigma}\,,$$

we obtain the equation

$$\frac{\partial}{\partial x_\alpha}\left(g^{\mu\nu}_{\sigma}\frac{\partial H}{\partial g^{\mu\nu}_{\alpha}}\right) - \frac{\partial H}{\partial x_\sigma} = 0$$

or[1]

$$\left.\begin{aligned} \frac{\partial t^{\alpha}_{\sigma}}{\partial x_\alpha} &= 0 \\ -2\kappa \cdot t^{\alpha}_{\sigma} &= g^{\mu\nu}_{\sigma}\frac{\partial H}{\partial g^{\mu\nu}_{\alpha}} - \delta^{\alpha}_{\sigma}H\,, \end{aligned}\right\} \tag{16.4}$$

where, on account of Equation 16.2, the second equation of Equation 15.1, and Equation 12.7

$$\kappa \cdot t^{\alpha}_{\sigma} = \frac{1}{2}\cdot\delta^{\alpha}_{\sigma}g^{\mu\nu}\Gamma^{\lambda}_{\mu\beta}\Gamma^{\beta}_{\nu\lambda} - g^{\mu\nu}\Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\sigma}\,. \tag{16.5}$$

It is to be noticed that t^{α}_{σ} is not a tensor; on the other hand Equation 16.4 applies to all systems of co-ordinates for which $\sqrt{-g} = 1$. This equation expresses the law of conservation of momentum and of energy for the gravitational field. Actually the integration of this equation over a *three-dimensional* volume $\mathscr{V}$ yields the four

[1]The reason for the introduction of the factor -2κ will be apparent later.

equations

$$\frac{\mathrm{d}}{\mathrm{d}x_4}\left\{\int t^4_\sigma \,\mathrm{d}\mathscr{V}\right\} = \int \left(lt^1_\sigma + mt^2_\sigma + nt^3_\sigma\right)\mathrm{d}S\,, \tag{16.6}$$

where l, m, n denote the direction-cosines of direction of the inward drawn normal at the element $\mathrm{d}S$ of the bounding surface (in the sense of EUCLIDean geometry). We recognize in this the expression of the laws of conservation in their usual form. The quantities t^α_σ we call the "energy components" of the gravitational field.
I will now give Equation 15.1 in a third form, which is particularly useful for a vivid grasp of our subject. By multiplication of the field Equation 15.1 by $g^{\nu\sigma}$ these are obtained in the "mixed" form. Note that

$$g^{\nu\sigma}\frac{\partial\Gamma^{\ \alpha}_{\mu\nu}}{\partial x_\alpha} = \frac{\partial}{\partial x_\alpha}\left(g^{\nu\sigma}\Gamma^{\ \alpha}_{\mu\nu}\right) - \frac{\partial g^{\nu\sigma}}{\partial x_\alpha}\Gamma^{\ \alpha}_{\mu\nu}\,,$$

which quantity, by reason of Equation 12.7, is equal to

$$\frac{\partial}{\partial x_\alpha}\left(g^{\nu\sigma}\Gamma^{\ \alpha}_{\mu\nu}\right) - g^{\nu\beta}\Gamma^{\ \sigma}_{\alpha\beta}\Gamma^{\ \alpha}_{\mu\nu} - g^{\sigma\beta}\Gamma^{\ \nu}_{\beta\alpha}\Gamma^{\ \alpha}_{\mu\nu}\,,$$

or (with different symbols for the summation indices)

$$\frac{\partial}{\partial x_\alpha}\left(g^{\sigma\beta}\Gamma^{\ \alpha}_{\mu\beta}\right) - g^{\gamma\delta}\Gamma^{\ \sigma}_{\gamma\beta}\Gamma^{\ \beta}_{\delta\mu} - g^{\nu\sigma}\Gamma^{\ \alpha}_{\mu\beta}\Gamma^{\ \beta}_{\nu\alpha}\,.$$

The third term of this expression cancels with the one arising from the second term of the field Equation 15.1; using Equation 16.5, the second term may be written

$$\kappa\cdot\left(t_\mu^{\ \sigma} - \frac{1}{2}\cdot\delta_\mu^{\ \sigma}t\right)\,,$$

where $t = t_\alpha^{\ \alpha}$. Thus instead of Equation 15.1 we obtain

$$\left.\begin{aligned}\frac{\partial}{\partial x_\alpha}\left(g^{\sigma\beta}\Gamma^{\ \alpha}_{\mu\beta}\right) &= -\kappa\cdot\left(t_\mu^{\ \sigma} - \frac{1}{2}\cdot\delta_\mu^{\ \sigma}t\right)\\ \sqrt{-g} &= 1\,.\end{aligned}\right\} \tag{16.7}$$

17. The General Form of the Field Equations of Gravitation

The field equations for matter-free space formulated in chapter 16 are to be compared with the field equation

$$\nabla^2\phi = 0$$

of NEWTON's theory. We require the equation corresponding to POISSON's equation

$$\nabla^2\phi = 4\pi\kappa\rho\,,$$

where ρ denotes the density of matter.
The special theory of relativity has led to the conclusion that inert mass is nothing more or less than energy, which finds its complete mathematical expression in a symmetrical tensor of second rank, the energy-tensor. Thus in the general theory of relativity we must introduce a corresponding energy-tensor of matter T_σ^α, which, like the energy-components t_σ (Equation 16.4 and Equation 16.5) of the gravitational field, will have mixed character, but will pertain to a symmetrical covariant tensor.[1]
The system of Equation 16.7 shows how this energy-tensor (corresponding to the density ρ in POISSON's equation) is to be introduced into the field equations of gravitation. For if we consider a complete system (e.g. the solar system), the total mass of the system, and therefore its total gravitating action as well, will depend on the total energy of the system, and therefore on the ponderable energy together with the gravitational energy. This will allow itself to be expressed by introducing into Equation 16.7, in place of the energy-components of the gravitational field alone, the sums $t_\mu^\sigma + T_\mu^\sigma$ of the energy-components of matter and of gravitational field. Thus instead of Equation 16.7 we obtain the tensor equation

$$\left.\begin{aligned}\frac{\partial}{\partial x_\alpha}\left(g^{\sigma\beta}\Gamma_{\mu\beta}^{\ \alpha}\right) &= -\kappa\cdot\left[\left({t_\mu}^\sigma + {T_\mu}^\sigma\right) - \frac{1}{2}{\delta_\mu}^\sigma(t+T)\right]\\ \sqrt{-g} &= 1\,,\end{aligned}\right\}\qquad(17.1)$$

where we have set $T = {T_\mu}^\mu$ (LAUE's scalar). These are the required general field equations of gravitation in mixed form. Working back from these, we have in place

[1] $g_{\alpha\tau}{T_\sigma}^\alpha = T_{\sigma\tau}$ und $g^{\sigma\beta}{T_\sigma}^\alpha = T^{\alpha\beta}$ are to be symmetrical tensors.

of Equation 15.1

$$\left.\begin{aligned} \frac{\partial \Gamma^{\alpha}_{\mu\nu}}{\partial x_\alpha} + \Gamma^{\alpha}_{\mu\beta}\Gamma^{\beta}_{\nu\alpha} = -\kappa \cdot \left(T_{\mu\nu} - \frac{1}{2} \cdot g_{\mu\nu} T\right) \\ \sqrt{-g} = 1\,. \end{aligned}\right\} \tag{17.2}$$

It must be admitted that this introduction of the energy-tensor of matter is not justified by the relativity postulate alone. For this reason we have here deduced it from the requirement that the energy of the gravitational field shall act gravitationally in the same way as any other kind of energy. But the strongest reason for the choice of these equations lies in their consequence, that the equations of conservation of momentum and energy, corresponding exactly to Equation 16.4 and Equation 16.6, hold good for the components of the total energy. This will be shown in chapter 18.

18. The Laws of Conservation in the General Case

Equation 17.1 may readily be transformed so that the second term on the right-hand side vanishes. Contract Equation 17.1 with respect to the indices μ and σ, and after multiplying the resulting equation by $\frac{1}{2} \cdot \delta_\mu^{\ \sigma}$, subtract it from Equation 17.1. This gives

$$\frac{\partial}{\partial x_\alpha} \left(g^{\sigma\beta} \Gamma^{\ \alpha}_{\mu\beta} - \frac{1}{2} \cdot \delta_\mu^{\ \sigma} g^{\lambda\beta} \Gamma^{\ \alpha}_{\lambda\beta} \right) = -\kappa \cdot \left(t_\mu^{\ \sigma} + T_\mu^{\ \sigma} \right) . \tag{18.1}$$

On this equation we perform the operation $\frac{\partial}{\partial x_\sigma}$. We have

$$\begin{aligned} &\frac{\partial^2}{\partial x_\alpha \, \partial x_\sigma} \left(g^{\sigma\beta} \Gamma^{\ \alpha}_{\mu\beta} \right) \\ &= -\frac{1}{2} \cdot \frac{\partial^2}{\partial x_\alpha \, \partial x_\sigma} \left[g^{\sigma\beta} g^{\alpha\lambda} \cdot \left(\frac{\partial g_{\mu\lambda}}{\partial x_\beta} + \frac{\partial g_{\beta\lambda}}{\partial x_\mu} - \frac{\partial g_{\mu\beta}}{\partial x_\lambda} \right) \right] . \end{aligned}$$

The first and third terms of the round brackets yield contributions which cancel one another, as may be seen by interchanging, in the contribution of the third term, the summation indices α and σ on the one hand, and β and λ on the other. The second term may be re-modelled by Equation 12.4, so that we have

$$\frac{\partial^2}{\partial x_\alpha \, \partial x_\sigma} \left(g^{\sigma\beta} \Gamma^{\ \alpha}_{\mu\beta} \right) = \frac{1}{2} \cdot \frac{\partial^3 g^{\alpha\beta}}{\partial x_\alpha \, \partial x_\beta \, \partial x_\mu} . \tag{18.2}$$

The second term on the left-hand side of Equation 18.1 yields in the first place

$$-\frac{1}{2} \cdot \frac{\partial^2}{\partial x_\alpha \, \partial x_\mu} \left(g^{\lambda\beta} \Gamma^{\ \alpha}_{\lambda\beta} \right)$$

or

$$\frac{1}{4} \cdot \frac{\partial^2}{\partial x_\alpha \, \partial x_\mu} \left[g^{\lambda\beta} g^{\alpha\delta} \cdot \left(\frac{\partial g_{\delta\lambda}}{\partial x_\beta} + \frac{\partial g_{\delta\beta}}{\partial x_\lambda} - \frac{\partial g_{\lambda\beta}}{\partial x_\delta} \right) \right] .$$

With the choice of co-ordinates which we have made, the term deriving from the last term in round brackets disappears by reason of Equation 12.2. The other two may be combined, and together, by Equation 12.4, they give

$$-\frac{1}{2}\cdot\frac{\partial^3 g^{\alpha\beta}}{\partial x_\alpha\,\partial x_\beta\,\partial x_\mu},$$

so that in consideration of Equation 18.2, we have the identity

$$\frac{\partial^2}{\partial x_\alpha\,\partial x_\sigma}\left(g^{\sigma\beta}\Gamma^{\ \alpha}_{\mu\beta}-\frac{1}{2}\delta_\mu^{\ \sigma}g^{\lambda\beta}\Gamma^{\ \alpha}_{\lambda\beta}\right)\equiv 0 \tag{18.3}$$

From Equation 18.3 and Equation 18.1, it follows that

$$\frac{\partial\left(t_\mu^{\ \sigma}+T_\mu^{\ \sigma}\right)}{\partial x_\sigma}=0\,. \tag{18.4}$$

Thus it results from our field equations of gravitation that the laws of conservation of momentum and energy are satisfied. This may be seen most easily from the consideration which leads to Equation 16.6; except that here, instead of the energy components t_σ^μ of the gravitational field, we have to introduce the totality of the energy components of matter and gravitational field.

19. The Laws of Momentum and Energy for Matter, as a Consequence of the Field Equations

Multiplying Equation 17.2 by $\frac{\partial g^{\mu\nu}}{\partial x_\sigma}$, we obtain, by the method adopted in chapter 16, in view of the vanishing of

$$g_{\mu\nu} \cdot \frac{\partial g^{\mu\nu}}{\partial x_\sigma} \,,$$

the equation

$$\frac{\partial t_\sigma{}^\alpha}{\partial x_\alpha} + \frac{1}{2} \cdot \frac{\partial g^{\mu\nu}}{\partial x_\sigma} T_{\mu\nu} = 0 \,,$$

or, in view of Equation 18.4

$$\frac{\partial T_\sigma{}^\alpha}{\partial x_\alpha} + \frac{1}{2} \cdot \frac{\partial g^{\mu\nu}}{\partial x_\sigma} T_{\mu\nu} = 0 \,. \tag{19.1}$$

Comparison with Equation 12.16b shows that with the choice of system of co-ordinates which we have made, this equation predicates nothing more or less than the vanishing of divergence of the material energy-tensor. Physically, the occurrence of the second term on the left-hand side shows that laws of conservation of momentum and energy do not apply in the strict sense for matter alone, or else that they apply only when the $g^{\mu\nu}$ are constant, i.e. when the field intensities of gravitation vanish. This second term is an expression for momentum, and for energy, as transferred per unit of volume and time from the gravitational field to matter. This is brought out still more clearly by re-writing Equation 19.1 in the sense of Equation 12.15 as

$$\frac{\partial T_\sigma{}^\alpha}{\partial x_\alpha} = -\Gamma^\alpha_{\sigma\beta} \Gamma_\alpha{}^\beta \,. \tag{19.2}$$

The right side expresses the energetic effect of the gravitational field on matter. Thus the field equations of gravitation contain four conditions which govern the course of material phenomena. They give the equations of material phenomena completely, if the latter is capable of being characterized by four differential equations independent of one another.[1]

[1]On this question cf. H. Hilbert, Nachr. d. K. Gesellsch. d. Wiss. zu Göttingen, Math.-phys. Klasse, 1915, p, 3.

Part IV.

MATERIAL PHENOMENA

20. Introduction

The mathematical aids developed in Part II enable us forthwith to generalize the physical laws of matter (hydrodynamics, MAXWELL's electrodynamics), as they are formulated in the special theory of relativity, so that they will fit in with the general theory of relativity. When this is done, the general principle of relativity does not indeed afford us a further limitation of possibilities; but it makes us acquainted with the influence of the gravitational field on all processes, without our having to introduce any new hypothesis whatever.
Hence it comes about that it is not necessary to introduce definite assumptions as to the physical nature of matter (in the narrower sense). In particular it may remain an open question whether the theory of the electromagnetic field in conjunction with that of the gravitational field furnishes a sufficient basis for the theory of matter or not. The general postulate of relativity is unable on principle to tell us anything about this. It must remain to be seen, during the working out of the theory, whether electromagnetism and the doctrine of gravitation are able in collaboration to perform what the former by itself is unable to do.

21. Euler's Equations for a Frictionless Adiabatic Fluid

Let p and ρ be two scalars, the former of which we call the "pressure", the latter the "density" of a fluid; and let an equation subsist between them. Let the contravariant symmetrical tensor

$$T^{\alpha\beta} = -g^{\alpha\beta} \cdot p + \rho \cdot \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} \frac{\mathrm{d}x_\beta}{\mathrm{d}s} \tag{21.1}$$

be the contravariant energy-tensor of the fluid. To it belongs the covariant tensor

$$T_{\mu\nu} = -g_{\mu\nu} \cdot p + g_{\mu\alpha} \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} \cdot g_{\mu\beta} \frac{\mathrm{d}x_\beta}{\mathrm{d}s} \cdot \rho \,, \tag{21.2}$$

as well as the mixed tensor[1]

$$T_\sigma{}^\alpha = -\delta_\sigma{}^\alpha \cdot p + g_{\sigma\beta} \frac{\mathrm{d}x_\beta}{\mathrm{d}s} \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} \cdot \rho \,. \tag{21.3}$$

Inserting the right-hand side of Equation 21.3 in Equation 19.2, we obtain the EULERian hydrodynamic equations of the general theory of relativity. They give, in theory, a complete solution of the problem of motion, since the four equations Equation 19.2, together with the given equation between p and ρ, and the equation

$$g_{\alpha\beta} \frac{\mathrm{d}x_\alpha}{\mathrm{d}s} \frac{\mathrm{d}x_\beta}{\mathrm{d}s} = 1$$

are sufficient, $g_{\alpha\beta}$ being given, to define the six unknowns

$$p\,, \quad \rho\,, \quad \frac{\mathrm{d}x_1}{\mathrm{d}s}\,, \quad \frac{\mathrm{d}x_2}{\mathrm{d}s}\,, \quad \frac{\mathrm{d}x_3}{\mathrm{d}s}\,, \quad \frac{\mathrm{d}x_4}{\mathrm{d}s}\,.$$

If the $g_{\mu\nu}$ are also unknown, the Equation 17.2 are brought in. These are eleven equations for defining the ten functions $g_{\mu\nu}$, so that these functions appear over-defined. We must remember, however, that the Equation 19.2 are already contained in the equations Equation 17.2, so that the latter represent only seven independent equations. There is good reason for this lack of definition, in that the wide freedom

[1]For an observer using a system of reference in the sense of the special theory of relativity for an infinitely small region, and moving with it, the density of energy $T_4{}^4$ equals $\rho - p$. This gives the definition of ρ. Thus ρ is not constant for an incompressible fluid.

of the choice of co-ordinates causes the problem to remain mathematically undefined to such a degree that three of the functions of space may be chosen at will.[2]

[2]On the abandonment of the choice of co-ordinates with $g = -1$, there remain *four* functions of space with liberty of choice, corresponding to the four arbitrary functions at our disposal in the choice of co-ordinates.

22. Maxwell's Electromagnetic Field Equations for Free Space

Let ϕ_ν be the components of a covariant vector – the electromagnetic potential vector. From them we form, in accordance with Equation 12.10, the components $F_{\rho\sigma}$ of the covariant six-vector of the electromagnetic field, in accordance with the system of equations

$$F_{\rho\sigma} = \frac{\partial \phi_\rho}{\partial x_\sigma} - \frac{\partial \phi_\sigma}{\partial x_\rho}\,. \tag{22.1}$$

It follows from Equation 22.1 that the system of equations

$$\frac{\partial F_{\rho\sigma}}{\partial x_\tau} + \frac{\partial F_{\sigma\tau}}{\partial x_\rho} + \frac{\partial F_{\tau\rho}}{\partial x_\sigma} = 0 \tag{22.2}$$

is satisfied, its left side being, by Equation 12.11, an antisymmetrical tensor of the third rank. Equation 22.2 thus contains essentially four equations which are written out as follows:–

$$\left.\begin{aligned}
\frac{\partial F_{23}}{\partial x_4} + \frac{\partial F_{34}}{\partial x_2} + \frac{\partial F_{42}}{\partial x_3} &= 0 \\
\frac{\partial F_{34}}{\partial x_1} + \frac{\partial F_{41}}{\partial x_3} + \frac{\partial F_{13}}{\partial x_4} &= 0 \\
\frac{\partial F_{41}}{\partial x_2} + \frac{\partial F_{12}}{\partial x_4} + \frac{\partial F_{24}}{\partial x_1} &= 0 \\
\frac{\partial F_{12}}{\partial x_3} + \frac{\partial F_{23}}{\partial x_1} + \frac{\partial F_{31}}{\partial x_2} &= 0\,.
\end{aligned}\right\} \tag{22.3}$$

This system corresponds to the second of MAXWELL's systems of equations. We recognize this at once by setting

$$\left.\begin{aligned}
F_{23} &= H_x \\
F_{31} &= H_y \\
F_{12} &= H_z \\
F_{14} &= E_x \\
F_{24} &= E_y \\
F_{34} &= E_z\,.
\end{aligned}\right\} \tag{22.4}$$

Then in place of Equation 22.3 we may set, in the usual notation of three-dimensional vector analysis,

$$\left.\begin{aligned}\frac{\partial H}{\partial t} + \operatorname{curl} E = 0 \\ \operatorname{div} H = 0\,.\end{aligned}\right\} \tag{22.5}$$

We obtain MAXWELL's first system by generalizing the form given by MINKOWSKI. We introduce the contravariant six- vector associated with $F^{\alpha\beta}$

$$F^{\mu\nu} = g^{\mu\alpha} q^{\nu\beta} F_{\alpha\beta} \tag{22.6}$$

and also the contravariant vector J^μ of the density of the electric current. Then, taking Equation 12.14 into consideration, the following equations will be invariant for any substitution whose invariant is unity (in agreement with the chosen coordinates):–

$$\frac{\partial F^{\mu\nu}}{\partial x_\nu} = J^\mu\,. \tag{22.7}$$

Let

$$\left.\begin{aligned} F^{23} &= H'_x \\ F^{31} &= H'_y \\ F^{12} &= H'_z \\ F^{14} &= -E'_x \\ F^{24} &= -E'_y \\ F^{34} &= -E'_z\,. \end{aligned}\right\} \tag{22.8}$$

which quantities are equal to the quantities $H_x, \ldots E_z$ in the special case of the restricted theory of relativity; and in addition

$$\begin{aligned} J^1 &= j_x \\ J^2 &= j_y \\ J^3 &= j_z \\ J^4 &= \rho\,, \end{aligned}$$

we obtain in place of Equation 22.7

$$\left.\begin{aligned}\operatorname{curl} H' - \frac{\partial E'}{\partial t} = j \\ \operatorname{div} E' = \rho\,.\end{aligned}\right\} \tag{22.9}$$

The Equation 22.2, Equation 22.6, and Equation 22.7 thus form the generalization of MAXWELL's field equations for free space, with the convention which we have established with respect to the choice of co-ordinates.

The Energy-components of the Electromagnetic Field.– We form the inner product

$$\kappa_\sigma = F_{\sigma\mu} J^\mu \tag{22.10}$$

By Equation 22.4, its components, written in the three-dimensional manner, are

$$\left.\begin{aligned} \kappa_1 &= \rho \cdot E_x + [j, H]_x \\ &\cdots\cdots \\ &\cdots\cdots \\ \kappa_4 &= -(j, E)\,. \end{aligned}\right\} \tag{22.11}$$

κ_σ is a covariant vector, the components of which are equal to the negative momentum, or, respectively, the energy, which is transferred from the electric masses to the electromagnetic field per unit of time and volume. If the electric masses are free, that is, under the sole influence of the electromagnetic field, the covariant vector κ_σ will vanish.

To obtain the energy-components $T_\sigma^{\ \nu}$ of the electromagnetic field, we need only give to equation $\kappa_\sigma = 0$ the form of Equation 19.1. From Equation 22.7 and Equation 22.10 we have in the first place

$$\kappa_\sigma = F_{\sigma\mu} \frac{\partial F^{\mu\nu}}{\partial x_\nu} = \frac{\partial}{\partial x_\nu} \left(F_{\sigma\mu} F^{\mu\nu}\right) - F^{\mu\nu} \frac{\partial F_{\sigma\mu}}{\partial x_\nu}\,.$$

The second term of the right-hand side, by reason of Equation 22.2, permits the transformation

$$F^{\mu\nu} \frac{\partial F_{\sigma\mu}}{\partial x_\nu} = -\frac{1}{2} \cdot F^{\mu\nu} \frac{\partial F_{\mu\nu}}{\partial x_\sigma} = -\frac{1}{2} \cdot g^{\mu\alpha} g^{\nu\beta} \cdot F_{\alpha\beta} \frac{\partial F_{\mu\nu}}{\partial x_\sigma}\,,$$

which latter expression may, for reasons of symmetry, also be written in the form

$$-\frac{1}{4} \cdot \left[g^{\mu\alpha} g^{\nu\beta} \cdot F_{\alpha\beta} \frac{\partial F_{\mu\nu}}{\partial x_\sigma} + g^{\mu\alpha} g^{\nu\beta} \cdot \frac{\partial F_{\alpha\beta}}{\partial x_\sigma} F_{\mu\nu} \right]\,.$$

But for this we may set

$$-\frac{1}{4} \cdot \frac{\partial}{\partial x_\sigma} \left(g^{\mu\alpha} g^{\nu\beta} F_{\alpha\beta} F_{\mu\nu} \right) + \frac{1}{4} \cdot F_{\alpha\beta} F_{\mu\nu} \frac{\partial}{\partial x_\sigma} \left(g^{\mu\alpha} g^{\nu\beta} \right)\,.$$

The first of these terms is written more briefly

$$-\frac{1}{4} \cdot \frac{\partial}{\partial x_\sigma} \left(F^{\mu\nu} F_{\mu\nu} \right)\,;$$

the second, after the differentiation is carried out, and after some reduction, results in

$$-\frac{1}{2}\cdot F^{\mu\tau}F_{\mu\nu}g^{\nu\rho}\frac{\partial g_{\sigma\tau}}{\partial x_\sigma}\,.$$

Taking all three terms together we obtain the relation

$$\kappa_\sigma = \frac{\partial T_\sigma{}^\nu}{\partial x_\nu} - \frac{1}{2}\cdot g^{\tau\mu}\cdot\frac{\partial g_{\mu\nu}}{\partial x_\sigma}T_\tau{}^\nu\,, \tag{22.12}$$

where

$$T_\sigma{}^\nu = -F_{\sigma\alpha}F^{\nu\alpha} + \frac{1}{4}\cdot\delta_\sigma{}^\nu F_{\alpha\beta}F^{\alpha\beta}\,. \tag{22.13}$$

Equation 22.12, if κ_σ vanishes, is, on account of Equation 12.3, equivalent to Equation 19.1 or Equation 19.2 respectively. Therefore the $T_\sigma{}^\nu$ are the energy-components of the electromagnetic field. With the help of Equation 22.4 and Equation 22.8, it is easy to show that these energy-components of the electromagnetic field in the case of the special theory of relativity give the well-known Maxwell-Poynting expressions.

We have now deduced the general laws which are satisfied by the gravitational field and matter, by consistently using a system of co-ordinates for which $\sqrt{-g}=1$. We have thereby achieved a considerable simplification of formulæ and calculations, without failing to comply with the requirement of general covariance; for we have drawn our equations from generally covariant equations by specializing the system of co-ordinates.

Still the question is not without a formal interest, whether with a correspondingly generalized definition of the energy-components of the gravitational field and matter, even without specializing the system of co-ordinates, it is possible to formulate laws of conservation in the form of Equation 18.4, and field equations of gravitation of the same nature as Equation 17.1 or Equation 18.1, in such a manner that on the left we have a divergence (in the ordinary sense), and on the right the sum of the energy-components of matter and gravitation. I have found that in both cases this is actually so. But I do not think that the communication of my somewhat extensive reflexions on this subject would be worthwhile, because after all they do not give us anything that is materially new.

Part V.

DISCUSSIONS ON NEWTONS THEORY

23. Newton's Theory as a First Approximation

As has already been mentioned more than once, the special theory of relativity as a special case of the general theory is characterized by the $g_{\mu\nu}$ having the constant values according to Equation 4.4. From what has already been said, this means complete neglect of the effects of gravitation. We arrive at a closer approximation to reality by considering the case where the $g_{\mu\nu}$ differ from the values of Equation 4.4 by quantities which are small compared with 1, and neglecting small quantities of second and higher order. (First point of view of approximation.)
It is further to be assumed that in the space-time territory under consideration the $g_{\mu\nu}$ at spatial infinity, with a suitable choice of co-ordinates, tend toward the values according to Equation 4.4; i.e. we are considering gravitational fields which may be regarded as generated exclusively by matter in the finite region.
It might be thought that these approximations must lead us to Newton's theory. But to that end we still need to approximate the fundamental equations from a second point of view. We give our attention to the motion of a material point in accordance with the Equation 9.1. In the case of the special theory of relativity the components

$$\frac{\mathrm{d}x_1}{\mathrm{d}s}, \quad \frac{\mathrm{d}x_2}{\mathrm{d}s}, \quad \frac{\mathrm{d}x_3}{\mathrm{d}s}$$

may take on any values. This signifies that any velocity

$$\mathbf{v} = \sqrt{\left(\frac{\mathrm{d}x_1}{\mathrm{d}x_4}\right)^2 + \left(\frac{\mathrm{d}x_2}{\mathrm{d}x_4}\right)^2 + \left(\frac{\mathrm{d}x_3}{\mathrm{d}x_4}\right)^2}$$

may occur, which is less than the velocity of light *in vacuo*. If we restrict ourselves to the case which almost exclusively offers itself to our experience, of v being small as compared with the velocity of light, this denotes that the components

$$\frac{\mathrm{d}x_1}{\mathrm{d}s}, \quad \frac{\mathrm{d}x_2}{\mathrm{d}s}, \quad \frac{\mathrm{d}x_3}{\mathrm{d}s}$$

are to be treated as small quantities, while $\frac{\mathrm{d}x_4}{\mathrm{d}s}$, to the second order of small quantities, is equal to one. (Second point of view of approximation.)
Now we remark that from the first point of view of approximation the magnitudes $\Gamma_{\mu\nu}^{\tau}$ are all small magnitudes of at least the first order. A glance at Equation 14.2

thus shows that in this equation, from the second point of view of approximation, we have to consider only terms for which $\mu = \nu = 4$. Restricting ourselves to terms of lowest order we first obtain in place of Equation 14.2 the equations

$$\frac{\mathrm{d}^2 x_\tau}{\mathrm{d}t^2} = F_{44}^{\ \tau} \,,$$

where we have set $\mathrm{d}s = \mathrm{d}x_4 = \mathrm{d}t$; or with restriction to terms which from the first point of view of approximation are of first order:–

$$\frac{\mathrm{d}^2 x_\tau}{\mathrm{d}t^2} = \begin{bmatrix} 44 \\ \tau \end{bmatrix} \quad (\tau = 1, 2, 3)$$
$$\frac{\mathrm{d}^2 x_4}{\mathrm{d}t^2} = -\begin{bmatrix} 44 \\ 4 \end{bmatrix} .$$

If in addition we suppose the gravitational field to be a quasi-static field, by confining ourselves to the case where the motion of the matter generating the gravitational field is but slow (in comparison with the velocity of the propagation of light), we may neglect on the right-hand side differentiations with respect to the time in comparison with those with respect to the space co-ordinates, so that we have

$$\frac{\mathrm{d}^2 x_\tau}{\mathrm{d}t^2} = -\frac{1}{2} \cdot \frac{\partial g_{44}}{\partial x_\tau} \quad (\tau = 1, 2, 3) \,. \tag{23.1}$$

This is the equation of motion of the material point according to NEWTON's theory, in which $\frac{1}{2} g_{44}$ plays the part of the gravitational potential. What is remarkable in this result is that the component g_{44} of the fundamental tensor alone defines, to a first approximation, the motion of the material point.

We now turn to the field Equation 17.2. Here we have to take into consideration that the energy-tensor of "matter" is almost exclusively defined by the density of matter ρ in the narrower sense, i.e. by the second term of the right-hand side of Equation 21.1 (or, respectively, Equation 21.2 or Equation 21.3). If we form the approximation in question, all the components vanish with the one exception of

$$T_{44} = \rho = T \,.$$

On the left-hand side of Equation 17.2 the second term is a small quantity of second order; the first yields, to the approximation in question,

$$+\frac{\partial}{\partial x_1}\begin{bmatrix} \mu\ \nu \\ 1 \end{bmatrix} + \frac{\partial}{\partial x_2}\begin{bmatrix} \mu\ \nu \\ 2 \end{bmatrix} + \frac{\partial}{\partial x_3}\begin{bmatrix} \mu\ \nu \\ 3 \end{bmatrix} + \frac{\partial}{\partial x_4}\begin{bmatrix} \mu\ \nu \\ 4 \end{bmatrix} .$$

For $\mu = \nu = 4$, this gives, with the omission of terms differentiated with respect to time,

$$-\frac{1}{2} \cdot \left(\frac{\partial^2 g_{44}}{\partial x_1^2} + \frac{\partial^2 g_{44}}{\partial x_2^2} + \frac{\partial^2 g_{44}}{\partial x_3^2} \right) = -\frac{1}{2} \cdot \nabla^2 \; g_{44} \, .$$

The last of Equation 17.2 thus yields

$$\nabla^2 \; g_{44} = \kappa \cdot \rho \, . \tag{23.2}$$

The Equation 23.1 and Equation 23.2 together are equivalent to NEWTON's law of gravitation.
By Equation 23.1 and Equation 23.2 the expression for the gravitational potential becomes

$$-\frac{\kappa}{8\pi} \cdot \int \frac{\rho \, \mathrm{d}\tau}{r} \, , \tag{23.3}$$

while NEWTON's theory, with the unit of time which we have chosen, gives

$$-\frac{K}{c^2} \cdot \int \frac{\rho \, \mathrm{d}\tau}{r}$$

in which K denotes the constant 6.7×10^{-8}, usually called the constant of gravitation. By comparison we obtain

$$\kappa = \frac{8\pi K}{c^2} = 1.87 \times 10^{-27} \, . \tag{23.4}$$

24. Behaviour of Rods and Clocks in the Static Gravitational Field. Bending of Light-rays. Motion of the Perihelion of a Planetary Orbit

To arrive at NEWTON's theory as a first approximation we had to calculate only one component, g_{44}, of the ten $g_{\mu\nu}$ of the gravitational field, since this component alone enters into the first approximation, Equation 23.1, of the equation for the motion of the material point in the gravitational field. From this, however, it is already apparent that other components of the $g_{\mu\nu}$ must differ from the values given in Equation 4.4 by small quantities of the first order. This is required by the condition $g = -1$.

For a field-producing point mass at the origin of co-ordinates, we obtain, to the first approximation, the radially symmetrical solution

$$\left.\begin{aligned} g_{\rho\sigma} &= -\delta_{\rho\sigma} - \alpha \cdot \frac{x_\rho x_\sigma}{r^3} \quad (\rho \text{ and } \sigma \text{ between 1 and 3}) \\ g_{\rho 4} &= g_{4\rho} = 0 \quad (\rho \text{ between 1 and 3}) \\ g_{44} &= 1 - \frac{\alpha}{r}\,. \end{aligned}\right\} \tag{24.1}$$

where $\delta_{\rho\sigma}$ is 1 or 0, respectively, accordingly as $\rho = \sigma$ or $\rho \pm \sigma$, and r is the quantity

$$+\sqrt{x_1^2 + x_2^2 + x_3^2}\,.$$

On account of Equation 23.3

$$\alpha = \frac{\kappa M}{4\pi}\,, \tag{24.2}$$

if M denotes the field-producing mass. It is easy to verify that the field equations (outside the mass) are satisfied to the first order of small quantities.

We now examine the influence exerted by the field of the mass M upon the metrical properties of space. The relation

$$\mathrm{d}s^2 = g_{\mu\nu} \cdot \mathrm{d}x_\mu \, \mathrm{d}x_\nu$$

always holds between the "locally" (chapter 4) measured lengths and times $\mathrm{d}s$ on the one hand, and the differences of co-ordinates $\mathrm{d}x_\nu$ on the other hand.
For a unit-measure of length laid "parallel" to the axis of x, for example, we should have to set

$$\begin{aligned} \mathrm{d}s^2 &= -1 \\ \mathrm{d}x_2 &= \mathrm{d}x_3 = \mathrm{d}x_4 = 0\,, \end{aligned}$$

Therefore

$$g_{11}\,\mathrm{d}x_1^2 = -1\,.$$

If, in addition, the unit-measure lies on the axis of x, the first of Equation 24.1 gives

$$g_{11} = -\left(1 + \frac{\alpha}{r}\right)\,.$$

From these two relations it follows that, correct to a first order of small quantities,

$$\mathrm{d}x = 1 - \frac{\alpha}{2 \cdot r}\,.$$

The unit measuring-rod thus appears a little shortened in relation to the system of co-ordinates by the presence of the gravitational field, if the rod is laid along a radius.
In an analogous manner we obtain the length of coordinates in tangential direction if, for example, we set

$$\begin{aligned} \mathrm{d}s^2 &= -1 \\ \mathrm{d}x_1 &= \mathrm{d}x_3 = \mathrm{d}x_4 = 0 \\ x_1 &= r \\ x_2 &= x_3 = 0\,. \end{aligned}$$

The result is

$$-1 = g_{22}\,\mathrm{d}x_2^2 = -\,\mathrm{d}x_2^2\,. \tag{24.3}$$

With the tangential position, therefore, the gravitational field of the point of mass has no influence on the length of a rod.
Thus Euclidean geometry does not hold even to a first approximation in the gravitational field, if we wish to take one and the same rod, independently of its place and orientation, as a realization of the same interval; although, to be sure, a glance at Equation 24.2 and Equation 23.4 shows that the deviations to be expected are much too slight to be noticeable in measurements of the earth's surface.

Further, let us examine the rate of a unit clock, which is arranged to be at rest in a static gravitational field. Here we have for a clock period

$$\begin{aligned} \mathrm{d}s &= 1 \\ \mathrm{d}x_1 = \mathrm{d}x_2 &= \mathrm{d}x_3 = 0\,. \end{aligned}$$

Therefore

$$\begin{aligned} 1 &= g_{44} \cdot \mathrm{d}x_4^2 \\ \mathrm{d}x_4 &= \frac{1}{\sqrt{g_{44}}} = \frac{1}{\sqrt{1 + (g_{44} - 1)}} = 1 - \frac{g_{44} - 1}{2} \end{aligned}$$

or

$$\mathrm{d}x_4 = 1 + \frac{\kappa}{8\pi} \cdot \int \frac{\rho\,\mathrm{d}\tau}{r}\,. \tag{24.4}$$

Thus the clock goes more slowly if set up in the neighbourhood of ponderable masses. From this it follows that the spectral lines of light reaching us from the surface of large stars must appear displaced towards the red end of the spectrum.[1] We now examine the course of light-rays in the static gravitational field. By the special theory of relativity the velocity of light is given by the equation

$$-\,\mathrm{d}x_1^2 - \mathrm{d}x_2^2 - \mathrm{d}x_3^2 + \mathrm{d}x_4^2 = 0$$

and therefore by the general theory of relativity by the equation

$$\mathrm{d}s^2 = g_{\mu\nu} \cdot \mathrm{d}x_\mu\,\mathrm{d}x_\nu = 0\,. \tag{24.5}$$

If the direction, i.e. the ratio $\mathrm{d}x_1 : \mathrm{d}x_2 : \mathrm{d}x_3$, is given, Equation 24.5 gives the quantities

$$\frac{\mathrm{d}x_1}{\mathrm{d}x_4}\,, \quad \frac{\mathrm{d}x_2}{\mathrm{d}x_4}\,, \quad \frac{\mathrm{d}x_3}{\mathrm{d}x_4}$$

and accordingly the velocity

$$\sqrt{\left(\frac{\mathrm{d}x_1}{\mathrm{d}x_4}\right)^2 + \left(\frac{\mathrm{d}x_2}{\mathrm{d}x_4}\right)^2 + \left(\frac{\mathrm{d}x_3}{\mathrm{d}x_4}\right)^2} = \gamma\,,$$

defined in the sense of EUCLIDean geometry. We easily recognize that the course of the light-rays must be bent with regard to the system of co-ordinates, if the $g_{\mu\nu}$ are not constant. If n is a direction perpendicular to the propagation of light, the HUYGENS principle shows that the light-ray, envisaged in the plane (γ, n), has the

[1] According to E. FREUNDLICH, spectroscopical observations on fixed stars of certain types indicate the existence of an effect of this kind, but a crucial test of this consequence has not yet been made.

curvature $-\frac{\partial \gamma}{\partial n}$.
We examine the curvature undergone by a ray of light passing by a mass M at the distance Δ. If we choose the system of co-ordinates in agreement with the accompanying Figure 24.1, the total bending of the ray (calculated positively if concave towards the origin) is given in sufficient approximation by

$$B = \int_{-\infty}^{+\infty} \frac{\partial \gamma}{\partial x_1} \, \mathrm{d}x_2 \, ,$$

while Equation 24.5 and Equation 24.1 give

$$\gamma = \sqrt{-\frac{g_{44}}{g_{22}}} = 1 + \frac{\alpha}{2 \cdot r} \cdot \left(1 + \frac{x_2^2}{r^2}\right) \, .$$

Carrying out the calculation, this gives

$$B = \frac{2\alpha}{\Delta} = \frac{\kappa M}{2\pi\Delta} \, .$$

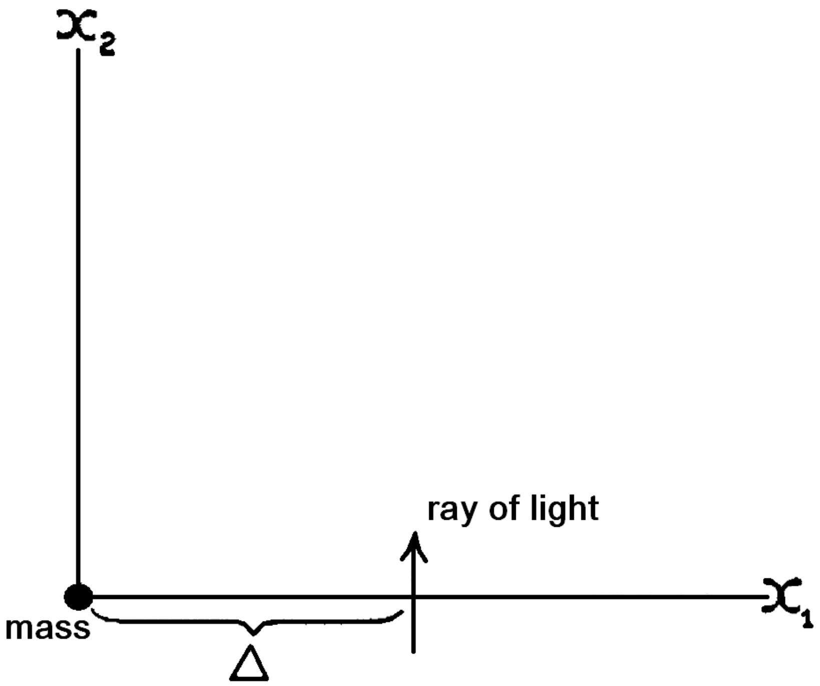


Figure 24.1.: Bending of a light ray by a mass M at distance Δ

According to this, a ray of light going past the sun undergoes a deflexion of 1.7″; and a ray going past the planet Jupiter a deflexion of about 0.2″.

If we calculate the gravitational field to a higher degree of approximation, and likewise with corresponding accuracy the orbital motion of a material point of relatively infinitely small mass, we find a deviation of the following kind from the KEPLER-NEWTON laws of planetary motion. The orbital ellipse of a planet undergoes a slow rotation, in the direction of motion, of amount

$$\epsilon = 24\pi^3 \cdot \frac{a^2}{T^2c^2 \cdot (1 - e^2)} \tag{24.6}$$

per revolution. In this formula a denotes the major semi-axis, c the velocity of light in the usual measurement, e the eccentricity, T the time of revolution in seconds.[2] Calculation gives for the planet Mercury a rotation of the orbit of 43″ per century, corresponding exactly to astronomical observation (LEVERRIER); for the astronomers have discovered in the motion of the perihelion of this planet, after allowing for disturbances by other planets, an inexplicable remainder of this magnitude.

[2]For the calculation I refer to the original papers: A. EINSTEIN, Sitzungsber. d. Preuss. Akad. d. Wiss., 1915, p. 831; K. SCHWARZSCHILD, *ibid.*, 1916, p. 189.

Part VI.

Appendix

25. Some Notes on Tensors

There exist two forms of representation for the components of a **four-vector**:[1]

$$\boxed{\begin{aligned} A^i &= \text{ contravariant} \\ A_i &= \text{ covariant} \end{aligned}}$$

The three spacial components of a four-vector A^i form a three-dimensional vector $\mathbf{A}$ regarding spatial rotation, i.e. a transformation that doesn't affect the time co-ordinate. The time component of a four-vector represents a three-dimensional scalar regarding these transformations. Therefore one can write for a four-vector

$$A^i = (A^0, \mathbf{A})\,.$$

Analogously for a **tensor** there exist three forms of representation for its components:

$$\boxed{\begin{aligned} A^{ik} &= \text{ contravariant} \\ A_{ik} &= \text{ covariant} \\ A^i{}_k &= \text{ mixed} \\ A_i{}^k &= \text{ mixed} \end{aligned}}$$

In the case of a mixed tensor there is in general to differentiate between the last two cases, depending on whether the first index is subscript and the second superscript or vice versa.

[1]See J. M. Lifschitz, L. D. Landau, Band II, Klassische Feldtheorie, 1997, p. 18.

The relation between the different types of the components is given through the following common rule:

> The sign of a component is changed through the raising or lowering of a spacial indices (1, 2, 3), but *not* by the raising or lowering of a time component (0).

This, e.g., gives:

$$\begin{aligned} A_{00} &= A^{00}; \quad A_{01} = -A^{01} \\ A_0^0 &= A^{00}; \quad A_0^1 = A^{01} \\ A_1^0 &= -A^{01}; \quad A_1^1 = -A^{11} \\ A_{11} &= A^{11}; \; \cdots . \end{aligned}$$